GHI Modeling Team Handbook

Originally written by **Chatham Borsh** in Spring 2017
Last Updated on May 15, 2021

Table of Contents

Current Modeling Tasks as of 4/10/2024

Hep C

Use https://iris.who.int/bitstream/handle/10665/357086/9789240052697-eng.pdf?sequence=1

9789240052697-eng.pdf (who.int)

Shen Copy of 2019 Model *New Treatment Guidelines* OFFICIAL Hep C Model from OR…

Treatment Guideline update

To fill in guidelines and make sure guidelines match up the efficacy.

- Have Ainsley check through Hep C efficacy data.

Malaria

- Check through proportions p.falc and p.vivax

Need to add in values for DALY>1 countries

Shen's Copy of 2019 Model Reconstruction of CS Automated template Malaria 2015 - Google Sheets

9789240015791-eng.pdf - Google Drive

TB

% TB incidence with known HIV status (R)

- Tell Sanya to automate this data

We can calculate the data of estimated % HIV status. Can be calculated from column O.

Where do we get the data for countries without data

- From L28

If the total global average isn't the same as the data then we need to calculate it by hand to see if it matches up.

- May need to propose updated formula
- Could potentially be automated

Google what are the treatments for XTR-TB

HIV

(old HIV model, just check Sanya's https://docs.google.com/spreadsheets/d/121bTpNnOmliNdH24mjjmKhPYe2WCQkUWaSxDKCwRhc4/edit?usp=drive_link)

Ainsley will explain to Sanya and verify her changes and check website): completed Cells BE, F, G18 to BE, F, G 27 Treatment Coverage By Region [Completed - Nicole looked at BE cols and they are wrong]

HepC)

- Done. Shen made it, Ainsley checked, Nicole verified everything and said: New adolescent treatment guidelines were confirmed by Ainsley. **Decide what to do about including them - Sanya should do it and Shen will send us the treatment guidelines by 15.4.24.**

(TB)

- Ainsley confirm everything is in the excel and says X made it, Y checked, Nicole verified everything cols A B C etc. then delete this point.
- E.g. L28: Implemented the global and have Shen double checked.
- Then for Col O: Percent TB cases with known TB status R draws on L28 if not automated and the formula is listed in R's comment and Ainsley downloaded the data and saw that the average for the countries with data is not the same as the global average for 2022 so **confirm with Sanya the formula makes sense and implement and her verify**. Ainsely asked Sanya to automate this data and indicated that this has to be done in the automation handbook. **When that is done double check both the automated data and formula it is input correctly before deleting this (and everything else) from the automation handbook.**
- L30: confirmed Shen checked.
- L31: Ainsley confirmed the definitions of pre- and xdr- and mdr- TB are the same as the WHOs. Shen double checked.

- L36: Find what proportion of MDR/RR have just MDR to break out estimates. Found and Shen checked
- L26 found Shen checked and L38: **could not find reverted to old data.**
- **BF and BI Ainsley ask Sanya to automate and put in automation handbook. We should probably proportion out the MDR% out of MDR/RR *ask Sanya where to do it so you don't mess up formulas and Shen double check. Double check all automated data.**
- **Note what you have on BD, BE, BG, BH in the excel sheet for future with full references Ainsley**. Nicole says: We will need to give this some thought as we will get database data where possible and necessary given the guidelines and then consider if we need to hand collect any data for these cols. Try to find what you need for the treatment guidelines this and we can talk about this on March 8.
- **Include the same data for Pre-XDR in col L as you have for MDR and XDR in case we need it for a model update where we have to implement separate cols for pre-XDR in our model if you have it. *Ask Sanya where to put it so as not to mess up the formulas. Have Shen double check the input data. Also include the datapoints for Pre/XDR where possible incase we have to bundle them.**
- **Link the email you sent in the automation handbook and here and follow up with Sanya until everything is done Ainsley.**
- **New treatment guidelines need to be implemented and values adjusted accordingly.**
-
- dalys*treatment coverage proprotion*efficacy/(1-treatment coverage proprotion*efficacy)
    - **figure out what in the data dictionary we need to calculate**
- After you answer the questions on the TB treatment guideline document and hand it to Shen with the sources to think about, you can look for effectiveness data (real world data) in the TB reports but if you cannot find it we will need you to look for some efficacy data with the team for all the treatments we need - start by reading the systematic review handbook and looking at older efficacy data if needed. We can talk about it at the next meeting. The other big task is to figure out what we need for each disease sub-type e.g. MDR/RR TB what are the col names that need to be inputed to do the relevant calculations. To figure this out do not just look at the current col names (though they with their old definitions) will be helpful as we just need the updated data for some disease subtypes. You should read the appendix of the most recent paper with TB in it in the tropical medicine journal on TB and walk through the model's calculations with that in hand so you can see how everything is done.

- We are basically looking for the cols that let us calculate the proportions of people who need treatment who are recieving it, the DALYs lost to the disease sub-type wil be based on the proportions with that sub-type if it is not in the IHME DALY database, and then there is the treatment impact col which does the calculations using those two data points and the efficacy data which I referenced above.

(Trachoma)

- Done. Jacleen made it, Ainsley checked, Nicole verified. Ainsley asked Sanya to automate col P and put the relevant info into the automation handbook. **Make sure Sanya does it.**

(NTD)

- Done. Jacleen made it, Ainsley checked, Nicole verified everything and all data points can be automated. Ainsley asked Sanya to automate col C-AM and put the relevant info into the automation handbook. **Make sure Sanya does it.**

(Malaria)

- Completed Shen added in values for DALY>1 countries, Ainsley checked the calculations.

(HIV)

- **Ainsley will detail some errors which Sanya needed to fix in an email and the modeling handbook, and put the details also in the excel sheet. Follow up as needed to fix Ainsley.**
- Shen verified data in cols 17-27 and rows BE-BG.

-
-

Overview by Chatham Borsh

I joined the project in Jan. 2016, and have spent 1½ years with the project. About 5-6 months in, the previous Excel leader, Junyi Dong, stepped away from GHI, at which point I began to learn enough to fill her role. At that point, there were several existing models. HIV, TB, and Malaria had models for 2010 and 2013, which were fairly accurate and had been the basis for the first iteration of GHI's website. There were also 2010/2013 models for newly-incorporated Neglected Tropical Diseases (NTD's): Schistosomiasis, Lymphatic Filariasis, and Soil-Transmitted Helminthiasis (Including Hookworm, Whipworm, and Roundworm). Onchocerciasis has a 2013 model only, due to lack of data for the earlier year. The NTD models were less polished than the website ones. Numerous bugs, errors, etc. were present across all of them.

Since then the Modeling team has made numerous changes to the Project. First, a great deal of time was spent searching through models, checking the numbers and isolating errors in code to fix. Next was combining individual models into a summary sheet. Originally, each disease model could be found in its respective folder inside ORS for Individual Disease. We need to eventually perform Sensitivity Analysis, which requires the replication of models. To replicate all at once, we combined each model into one spreadsheet, now titled "ORS Global Burden of Disease 2013 (2010/2013)". We also produced a second copy of the spreadsheet using DALY data from a more recent Global Burden of Disease Report, which contained models for 2010 and 2015. The 2015 models have been updated with 2015 data points, where applicable.

For each model, we have also implemented new "efficacy algorithms". The Systematic Review team has compiled huge lists of efficacy studies. The algorithm utilizes an AVERAGEIFS function to sort through those sheets based on several possible factors (year, dosage, whether the study was randomized, country, etc.) to determine the average efficacy for the desired country. Currently, the data provided by Systematic Review is fairly scarce; the algorithms only pull a very limited number of points, meaning that much of our model is calculated from the resulting regional/global averages.

We have also written preliminary Technical Guides, Powerpoints, and papers for each disease. Most still require some checking over/editing, and the Powerpoints need to be made more consistent. Also important are changes to the manufacturing sheets. We have added 2015 sheets to TB and HIV, as well as 2010, 2013, and 2015 ones for Malaria. More details on those in Important Sheets/Folders.

How to use this manual

Welcome to the GHI modeling team! Contained in this manual contains almost everything you need to know to be successful in this team. It is a living document, meaning that as the needs of the team develop, so does this document.

Your First Read-through

Your team leader will ask you to read this manual on joining the team; this may feel like a daunting task. A lot of the modeling team's work includes utilizing excel in ways students may never have done before. Therefore, it is natural to experience a learning curve. A lot of the information you read the first time will go over your head, and that is ok. Once you start working with the models, you will be able to understand it better, and actually utilize the information.

That being said, the more information you can retain now, the more it will help you later. The most important sections to focus on are those that go over how the models work. Focus mainly on the sections: rules and good practices, what is GHI?, excel formulas worth knowing, and individual disease model explanations.

The other sections, while important and helpful, may not apply to the work you do in the early weeks of your internship. Look over what information is in these sections, and know you can refer back to them if needed in the future. Know that this manual is your first resource if you do not know how to do something; you can refer to it before emailing your team leader for clarification.

Your team leader may also assign you the Modeling Handbook Training Assignment. This is a good resource for learning about the different sections in the handbook.

For Experienced Members

Once you have gone through your training, and read the manual the first time, this does not mark the end of your need for it. You can continue to refer back to it when you need help.

You may need to refer to sections you never had to before. Like the For Team Leaders section at the end of the manual. This section is only really relevant to team leaders, but is important to refer to if considering this role, or when you actually take it on. It contains a collection of knowledge and advice from previous leaders that can help you in your time as one.

Sometimes, the manual might not have all of the information you may need. You may have to figure something out on your own. The needs of the team are always evolving, and the projects we work on could be very different to what we have done in the past. In these cases, it is important to add to the manual any training assignments or explanations you have made for this process.

What is GHI?

Chatham Borsh introduced the work of the modeling team in his overview. Since he was leader, there have been many changes in the team and models, but the overall goal remains the same: to create models to calculate the impact of pharmaceuticals in different countries and identify areas in need of greater access.

Our models, while regularly updated, also follow the same general design, which you will learn as you work on this team. These models contain the information on the GHI website. *When you read individual disease model explanations and website checking sections of the manual, you will learn where the information is located within the ORS.*

The website and intranet give a good overview of important resources you will learn about in this manual, as well as the current goals of the team.

Professor Hassoun gives a good overview of the project in the UN high-level report. Take the time now to read through the paper before moving on to later sections. Take care to learn about the impact score, and how it is calculated.

Rules and Good Practices

Before you get started on the team, it is important to know some general rules and practices we all must follow to ensure we can all collaborate most effectively. Here is a list of some general rules and practices you should be aware of.

****You must abide by these rules in order to remain a member of the team. Failure to do so can result in loss of important time and materials, leading to the postponing/loss of publications, which wastes not only your time, but also that of the entire team.****

- Reference your work
    - When you get further along in your time here at GHI, you will be tasked with creating models, or researching guidelines for new models. Before you do anything, it is important to make sure all sources are clearly cited with both a link to the source as well as a google drive link. If we do not have this information, all of your work could end up useless, or cost a lot of time locating this information. We can not submit a model to a journal if it is not properly referenced. See the referencing section of the manual for further information on how to do so.
- Keep proper folder organization
    - As you gain access to the rating system folder, you will be required to ensure all of your work is properly organized and all documents are tracked in this document. If you are unsure which folder you should put certain work in, contact your team leader and/or Professor Hassoun.
- Transfer all ownership to Professor Hassoun

    - As you start adding to the rating system folder, it is important to make sure the ownership of all documents you have created has been transferred to Professor Hassoun's Binghamton account. This ensures that after you graduate, your work is still accessible to new members, even if your Binhanton email has been deleted.
- Respond to emails within 24 hours
    - As you will learn, we are a highly collaborative team. As you get more involved, other team members will reach out to you looking for help with assignments. It is important to respond quickly to let them know when you can get it done, or if they should contact someone else to help them.
- If you do not understand something—ask!
    - We want every member to use their time as efficiently and effectively as possible. It is ok to not understand something right away. If you have spent 0.5-1 hour trying to understand something, and you still do not get it, reach out to your team leader or another senior member for help. If they do not respond in 24 hours, reach out to Professor Hassoun. In the time you wait for a response, you can move on to other projects.
- Keep track of your assignments
    - You will be thrown a lot of work seemingly all at once. It is your responsibility to keep track of it in a way that works best for you. Many members like to keep a list where they write down all assignments they have been given. This ensures that nothing gets lost in the woodwork, even as your priorities shift between more urgent assignments.
- Link your work in the updates
    - When filling out your updates each week, it is important to link you work so your leaders know what you have been working on, and can see your progress

How to Use Google Sheets

The modeling team primarily uses google sheets for its work. Therefore, all modeling team members need to be familiar with the program. While many college students have used google sheets in their courses, few have used it as in depth as we do on the modeling team. While it may seem daunting at first, when you break it down, the formulas are a lot simpler than you think.

Before we get started learning about the commonly used formulas in GHI, it is good to watch these introductory videos to get familiar with the program: beginners video and full tutorial(start at 13:14).

Utilizing features like autofill, and keyboard shortcuts, like the F4 key in absolute cell referencing, can make a big difference in the amount of time it takes to work on a specific section. Therefore, it is also important to explore the different shortcuts on your specific computer as well as the formulas themselves.

Excel Formulas Worth Knowing

- Commonly used functions – if unfamiliar, google the function name. support.office gives explanation of inputs, syntax, applications.
    - **AVERAGEIFS** – average of a range if corresponding ranges meet given criteria. Can be used to find average values for a given country/drug/year. SUMIFS/COUNTIFS can accomplish similar things.
    - **VLOOKUP** – searches the 1st column of a range for a specified value, returns corresponding value in another specified column in the range.
    - **IFERROR** – Precautionary function, an odd dataset/ lack of data for a given drug/region/whatever can result in an error message, which will be reflected in all sheets that call on that result. IFERROR can be used to return 0 in case of an error.
    - **IMPORTRANGE** – Uses a URL to copy a range from a different spreadsheet. Fairly effective, though a bit slow to update and requires some linking/access approval.
- Common GHI formulas – Formulas that we use across multiple models, which occur in locations and patterns that should become recognizable and predictable
    - **(X*Y*Z)/(1-(Y*Z))** – our impact formula. Checking what columns are referenced should give DALYs, efficacy, and treatment coverage. Should be found in Impact columns.
        - Impact score formula: (DALYs*Treatment Coverage*Efficacy)/(1-(Treatment Coverage*Efficacy))
    - **Nested IF/IFERROR/AVERAGEIFS Functions** – In formulas in which one or more of these functions are present or contained within each other, it is generally a way of implementing fallback data. If we lack specific, country-level data, we often then default to the WHO region average, and then the global average. Generally occurs in efficacy/treatment coverage columns.
- Commonly used formatting functions
    - SORT Range
        - Highlight the group of cells you'd like to sort.
        - Click Data and then Sort range.
        - If your columns have titles, click Data has header row.
        - Select the column you'd like to be sorted first and choose a sorting order.
        - To add another sorting rule, click Add another sort column.
        - Click Sort
    - FREEZE
        - Select a row or column you want to freeze or unfreeze.

        - At the top, click View and then Freeze.
        - Select how many rows or columns to freeze.
        - (To unfreeze, select a row or column. Then, at the top, click View and then Freeze and then No rows or No columns)

    - FILTER
        - Select a range of cells.
        - Click Data and then Create a filter.
        - To see filter options, go to the top of the range and click Filter.
            - Filter by condition: Choose conditions or write your own.
            - Filter by values: To hide data points, uncheck the box next to the data point and click OK.
            - Search: Search for data points by typing in the search box.
            - Filter by color: Choose which text or fill color to filter by. You can filter by conditional formatting colors, but not alternating colors.
        - To turn the filter off, click Data and then Turn off filter
- Now that you have been exposed to some of the excel formulas we used in GHI, it's time to practice using them. Here is a link to the Beginner Excel Training Assignment for Modelers. This assignment helps walk you through the formulas and practice using them.

Important Folders

Here listed are the important folders in the rating systems folder, as well as an explanation of what each subfolder contains. This is most important to more senior team members, as they have access to the rating systems folder. That being said, new members should note where this information is found, for when they gain access to the rating systems folder.

ORS 2022

contains all changes made in 2022. It follows a similar organization to the ORS 2020 and 2015 folders.

- Working Models
    - TB: Contains the new TB guidelines Lily found in the Spring 2022 semester
    - ORS: contains the most recent versions of the ORS 2022 (2010/2013) and the ORS 2022 (2015/2017)
    - Hep C: contains the changes made in the Summer by Lily and Sonia as well as the most recent copy of the Hep C paper
- Website Check Fall-Winter
    - contains the website check we did in the fall 2022-winter 2023
- NTD paper
    - contains the NTD paper and NTD supplement
    - ORS with NTD edits: contains edited versions of the ORS used for the NTD paper, the NTD graphs, and the calculations done for the paper

Automation project

- The automation project folder contains the information needed for the automation project, which is described in this manual.
- contains
    - working models: contains the output data, model outline, and blank model templates for each model we have
    - Automation handbook: contains the information needed for the modeling team to use the automation tool and find non-automated data

ORS 2020

Contains:

- Working Models
    - This folder contains the 2017 models and their associated data for the diseases HIV, TB, NTDs, Malaria, Hep C (Zeke's original model/writeup and Alyssa's edited model/writeup) and Trachoma
    - This folder also contains all data and rebuilt models for the years 2010, 2013, and 2015 that we worked on revamping during the 2021 Spring semester
- Archive
    - Contains documents that are no longer used
- Coronavirus
    - Research and model ideas on Coronavirus
- Sources Information
    - Source impact summary stored here
    - Newest sources paper writeup is located here
- Meeting Minutes
    - Contains the meeting minutes/agendas for each semester
- Model Outlines
    - General disease instructions on how to build a future model
- ORS Copies
    - These are various uptodate ORS copies that are used by modeling members
- Planning Ahead
    - Documentation on what to do in the following semesters
- Training Assignments
    - General documents/sheets that explain basic modeling tasks and assignments

- Website Necessities
    - Any documents or write ups that give data or an explanation for areas that have been updated on the website (company size research, forecasting tool, etc.)

ORS 2015

- Two google sheets, ORS Global Burden of Disease 2013 (2010/2013) and ORS Global Burden of Disease 2015 (2010/2015). Contain our models, efficacy studies, summary tabs.
- One google doc explaining the ORS 2015 subfolders
- ORS Errors
    - Recorded Errors and fixes on the ORS
- Cost Effectiveness Data
    - Contains the cost-effectiveness of drugs by manufacturer and disease
- Revenue Data
    - Contains the revenue data for the patent holders
- Training Videos
    - Contains videos that describe the models
- Possible New Models
    - Initial work on possible 2012 and 2016 NTD models
    - Initial ideas for future models
- Data Sources
    - Contains documents that provide data for the models and instructions on how to retrieve the data
- Archive
    - Contains documents that are no longer used
- Resources for the Papers
    - Contains documents that are used in the papers
- Technical Documents
    - Contains documents (presentations, docs, and PDFs used on the website) that explain how the models work
- Reports for Website
    - Contains company reports for website and all documents used to create it
- Treatment and Clinical Guidelines
    - Contains references and an archive folder
- Papers
    - Contains academic papers submitted or being prepared for submission
- Sensitivity Analysis
    - Contains the SA tests run to check the assumptions in the given ORS

- Manufacturing Company Evaluations
    - Contains data for the manufacturing model
- ORS Copies for Modeling (old)
    - These are old copies of ORS that new modeling members would use
- Source Implementation Issue
    - A recording of the source implementation issue resolve

ORS 2014 (Website)

Contains old website model (Version 63). Useful for looking at old iterations of model sheet. If a formula has a referencing error, or is missing a citation/explanation, check the older version of the model.

Important Models

Here listed are the important models for the GHI team. Only group leaders are allowed edit access to our Online Rating System (ORS). This is because it contains the information present on our website, and any changes to the document could lead to errors in our website. Modeling team members should be given access to drafts of these models for their work. Make sure to ask your group leader for the most recent version of the model every time you are given a task, as the ORS is constantly updated to ensure it is most accurately presenting the data calculated.

ORS 2022 (2010/2013)

- Most recent version of the ORS 2010/2013
- edited by Lily in the Fall 2022 semester
- present in the ORS 2022 folder
- minor errors in the DALY and Need tabs were fixed and documented in the ORS Summer/Fall edits/errors explanation document

ORS 2022 (2015/2017)

- Most recent version of the ORS 2015/2017
- edited by Lily and Sonia in the Summer and Fall 2022 semesters
- present in the ORS 2022 folders
- major edits:
    - listed in the ORS Summer/Fall edits/errors explanation document
    - addition of Hep C tab
    - integration of Hep C into company tab
    - deletion of 2010 data (2010 data in the ORS 2010/2013 were the correct versions)

- minor formatting changes to make it easier for the CS and modeling team to read and understand

Old versions of the models:

ORS Global Burden of Disease 2013 (2010/2013)

- Originally titled Impact Summary (Junyi)
- Contains Impact Score models for each disease we study.
- DALY data taken from Global Burden of Disease study 2013.
- TABS:
    - Company – main summary tab. Provides a live-updating summation of all our calculations in a format that the website can call upon. Any alterations in layout (adding or moving rows, columns) must be noted to web designer.
        - For both 2010 and 2013, lists each drug studied in our model, grouped by patent-holder. Below are total impact scores for the drug, separated by disease [cells pull from Impact Score tab]. Col AR-AU sum Impact, DALYs, and Need. Col AW on is the 2013 section.
        - @ Cell C26 – Manufacturing table. Pulls info from Manufacturing tabs, summarizes.
        - @ Cell A91 – summaries of efficacy/treatment coverage. Green-header table: broken down by sub disease. Orange table: final values for overall diseases, weighted if necessary by sub diseases portion of total Need .
    - Impact Score – Draws country level Impact Scores for each drug from the models.
    - DALY – Country level DALY data
    - NEED – Sum of Impact Score and DALY tabs; Theoretical DALYs in the absence of any treatment.
    - Individual Disease Models – Generally each have a 2010 model, 2013 model, and efficacy tab. More on these to follow.
    - Manufacturing Tabs – Integrate results of manufacturing sheets. Percentages based on GPRM data should not change after calculfmation, so full datasets unnecessary to include. Pulls Impact Scores from Company Tab, and returns results to Table at Cell C26.

ORS Global Burden of Disease 2015 (2010/2015)

- Copy of ORS Global Burden of Disease 2013 (2010/2013), with 2015 replacing 2013
- DALY data taken from GBD study for 2015, other 2015 data updated.
- Functionally identical to the other ORS, although the Company Tab has extra charts/graphs to the right of the company scores.

Individual Disease Model Explanations

Now that you know where the information is found in our folders, it is important to understand what is present in the ORS. The heart of the modeling team is the ORS, so all of the folders center around it.

Take a look at your copy of the ORS. Note that it is made up of multiple tabs. The first few tabs are used by the CS team to put the information on the website. They contain summaries of the impact of different drug regimens from different companies (Company tab), or the impact for each regimen for each country (Impact score tab). This information is all pulled from the individual disease models.

The disease models all serve to calculate the impact scores of the drug regimens used to treat said illness. Most models progress the same way from left to right:

- DALYs
- Treatment Coverage
- Efficacy (Country level pulled via algorithms, then estimated by region/globally)
- Impact
- Due to the specifics of each disease, variation will occur.

Each disease has a separate efficacy tab located at the end of the list of tabs. Each contains the list of studies generated by the Systematic Review team that contain the efficacy data. These tabs pull from that group's sheets via an ImportRange function, so we don't make direct alterations to them. If some formatting needs to be altered, contact the SR team.

Here are the resources available for understanding each of the models individually. Be sure to watch the videos and read the technical documents carefully and refer back to them as needed.

Malaria

Overview: The Malaria model is one of the easier models for new members to understand. If you understand the Malaria model, it is easier to understand the other models. Therefore, make sure to take your time to go over how this model works before going on to the next.

Note that the model is Split between two strains, *p. vivax* and *p. falc*. The latter is more prominent. Columns E/F give us treatment coverage. Col F, ACT treatment, is used as treatment coverage for *p.*

falc. E is all treatments, so the difference between the two, E-F, is treatment coverage for *p. vivax*. % *p. vivax* and % *p. falc* columns are multiplied by total DALYs to give the sub-disease DALYs.

2017 Malaria Model Folder

Malaria 2015 overview video**:** This video explains the different columns of the models and the purpose they serve.

The Malaria technical document and powerpoint is also a good resource for understanding the model.

How we build the model: Malaria Model Instructions

HIV

Overview: Anti-Retroviral drugs are always used in combination, making HIV more complicated than most other models. Each combination has a different efficacy, and is used to varying degrees among different populations. Populations receiving treatment are grouped by age (Adult vs Child), geographic group (A vs B, Non-Americas vs Americas), and line of treatment (1^{st} vs 2^{nd}). Score for a drug is the sum of the scores of each of the regimens containing that drug, weighted by the percentage of DALYs receiving the given regimen and using each regimen's specific efficacy.

Antiretroviral therapy occurs over a patient's lifetime. As such, we estimate one year's impact by dividing by the estimated lifespan of a patient. There is a maximum of 35 years, with possibly lower time spans by country. The final impact score for a country is divided by (100/ (100 - Retention rate in column K)). That value gives expected lifespan in years. For example, 100/ (100 -97.14) = 34.96 years.

2017 HIV Model Folder

HIV 2015 Overview video

HIV presentation

How we build the model: HIV Model Instructions

TB

Overview: The TB model is our most complex model. Note how there are four main types of TB we investigate in our model. This makes our model particularly complex. Be sure to refer to the powerpoint and explanation video to understand the model.

2017 TB Model Folder

TB explanation video

TB powerpoint

How we build the model: General TB Model Instructions, 2017 Model Instructions

NTDs

Overview: We group together 6 diseases here which are all classified as Neglected Tropical Diseases (NTDs). These are: Lymphatic Filariasis (LF), Schistosomiasis (Schist), Whipworm, Hookworm, Roundworm, Onchocerciasis (Onch). This model is probably most simple due to the fact that most NTD's are easily treatable with antibiotics and/or antiparisitics.

LF

Simple model. Treatment coverage derived from PCT databank.

Onchocerciasis

Simple model.

Schistosomiasis

Treatment Type, i.e. Drug Used, is determined by which other diseases are endemic to the area. A flowchart in the methodology explains which treatments are used for what combination of endemic diseases. Under the ONCHO, LF, SCH, and STH Level columns, a 1 indicates endemicity. The rest of the model is straightforward.

STH

Contains 3 sub-diseases, Roundworm, Hookworm, and Whipworm. DALYs and Treatment coverage are split between School Age Children (SAC), from 5-14 years old, and Pre-SAC, from 1-4 years old.

The 2013 STH model does not have specific DALY data for the 3 sub-diseases, only overall STH DALYs. The proportions of the 2010 model (e.g. 30% whipworm, 23% hookworm, 47% roundworm) were used to proportion each country's 2013 DALY data between them

NTD overview video

NTD presentation

Hep C

(Has its own training video)
(New training video)

Overview: Since Hep C has six different genotypes, this model contains an added layer of complexity. Col F containing treatment coverage, notice that this only currently uses regional data. With columns

G through L, this will show the percentage of each genotype within each country. Once broken into these six separate cases, which are summed up at the end to find the total impact score, each efficacy and impact score is calculated.

Recent Updates: During the Spring 2021 semester, it was recognized that the treatment regimens used in the original Hep C Model may not be accurate for the 2015 model year, as well as the fact that it was a very long list of regimens. For this reason, Alyssa was tasked with trying to find the best list of shortened treatment guidelines to implement into the model, then implement this new information into the model. Methodology of the new shortened regimen list was included on the doc and the newest form of the written protocol reflects these changes. If this new model is seen as more accurate, tech docs/powerpoints must be completed and a new or updated Hep C video may need to be made.

Hep C overview video

Trachoma

Overview: The Trachoma model is the simplest model in the ORS. This is because there is only one drug used to treat the disease.

Trachoma Overview Video

Other Resources

Some other resources for learning how to use the models are the PLOS paper and TMIH paper. Training assignments have been made for both papers, which your team leader will likely assign to help you go through the calculation process of the paper.

PLOS paper training assignment

TMIH paper training assignment

Manufacturing/Source Sheets

Training Video **(notice both videos (Sources and Manufacturers) pertain to this section)**

Here is the link to the new source explanation video

Here is the latest explanation

Overview: Our models provide Impact Scores for each drug. We attribute those scores to patent holders, but we can also attribute them to the drug manufacturers. Each disease necessitates its own separate manufacturing sheet/calculations. Simply put, if a company produced/shipped 20% of Lamivudine [3TC] in 2010, they can receive 20% of 3TC's impact calculated through our models. Percentages for each manufacturer for a drug are calculated from WHO Global Price Reporting

Mechanism and Daily Dose info. Impact Scores are imported from models, and distributed by those percentages, giving each manufacturer's impact score. This is then displayed on the GHI Website.

Information on how to set up new Manufacturing Sheets: Manufacturing Sheet folder

- How to calculate Cost-Effectiveness: **Instructions**
- Data taken from WHO Global Price Reporting Mechanism: Gives information on individual shipments of drugs (#units per package, location, Strength, Manufacturer)
- Sheets generally contain 1000s of rows of data, followed by our actual calculations.
    - E.g, in HIV 2013, rows 1-12801 are data from GPRM, while the following 200 rows contain important code.
- A column added to the data set calculates Lives Saved over a year by a given shipment of drugs: Lives = TNU/(DD*365), where TNU = total number of units, DD = Daily Dose.
    - 50,000 units of 3TC with a daily dose of 2 units would give ~68.5 lives saved.
        - 50,000/(2*365)
- Daily Dose info derived from company websites/multiple sources.

  The following is all calculated below the dataset:

- Those lives saved are summed for a given drug for each company that manufactured that drug (Column C). e.g. Aspen Pharmacare produced 3TC in 2010 that saved 98548.93 lives.
- Lives saved by all regimens containing the drug are also added by company. (Col D) E.g. Total lives saved by regimens produced by Aspen of Stavudine + Lamivudine [d4T+3TC], Stavudine + Lamivudine + Nevirapine [d4T+3TC+NVP], Zidovudine + Lamivudine [ZDV+3TC], and Zidovudine + Lamivudine + Nevirapine [ZVD+3TC+NVP] = 65548.44. Added to original 3TC lives saved, Sum of All Regimens = 98548.93 + 65548.44 = 164097.37 lives saved by 3TC produced by Aspen Pharmacare.
- Summing these numbers across all companies gives total lives saved for the drug, individually and in regimens. Each company's lives saved is a percentage of this number (Col E), which is the percentage of Impact DALY's they are then attributed (Col F).
- Company Size Data is also very important to this element of modeling since it correlates the company's revenue to its overall impact. During the Spring 2021 semester, we updated this data collection so that the most updated data could be implemented into the website. Company size data collection includes finding the companies yearly revenue, relating it to the company impact, and then determining the relationship between larger/smaller companies and their ability to alleviate disease burden. We find that including sales from alliance companies must be taken into consideration when calculating a company's size because it impacts the company's yearly revenue.

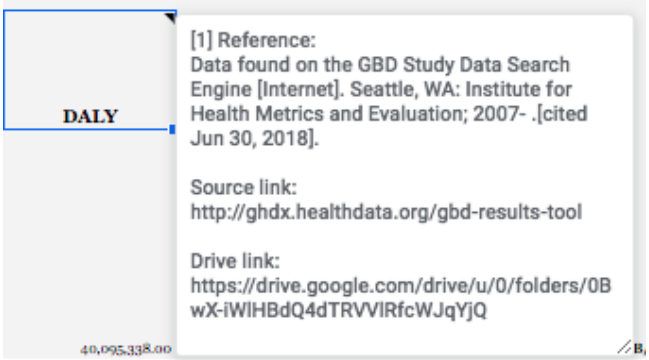
- The Source paper is a good resource for understanding how the source model works

Referencing

It is extremely important that you create a reference for all data collected as soon as you collect it. Any data that is implemented to the model without a reference will not be considered acceptable for use. Therefore, before you use any data you have found, it is important to properly reference it, as well as create a google drive link of the source.

Our references contain three parts: the actual reference, the link to the source, and the link to the google doc with the raw data. We keep both the actual link and the google doc link in case the actual source link gets broken. To input the reference, use the [insert] note feature. Make sure to label data sheets with years because this data may change once databases are updated

Here is an example of what the final reference should look like:



The most common style for referencing is vancouver, as it usually contains all of the needed information. Different journals may ask for references to be formatted in different ways, therefore, it is important to include all of the necessary information so they can be altered later to the needed style.

For more information on referencing, see the "Referencing Handbook" from the general team.

In order to reinforce the importance of referencing to interns and train them to learn all forms of referencing, the GHI Referencing Task is a great introduction. The basic outline of the initial reference should include at least the author, year of publication, location of publication, publisher, editions or volumes as necessary, page numbers, name of document being referenced, and the larger container where the referenced document is held.

Data Collection

The data we use in our models comes from various sources. We have specific data banks you can see often cited throughout our models that provide information on DALYs, prevalence, and treatment coverate.

Several useful sites for our basic data needs:

- Global Burden of Disease Report http://ghdx.healthdata.org/gbd-results-tool
    - Releases a yearly study. Our source for country DALY and Prevalence numbers.
    - engage@healthdata.org. - Requests for data from the Institute for Health Metrics and Evaluation (IHME) that isn't readily available on their site
    - The "GBDx Collection process for DALY/Prevalence" training assignment demonstrates how we use this source to find information for each disease.
- WHO PCT Databank http://www.who.int/neglected_diseases/preventive_chemotherapy/databank/en/
    - Information on NTD's. Provides Treatment Coverage, Drug Used
- Global Price Reporting Mechanism http://apps.who.int/hiv/amds/price/hdd/ - Manufacturer's Info
- Some diseases require specific sites/documents
    - TB World Report, Global TB Database http://www.who.int/tb/country/en/ – Numbers on sub diseases of TB; MDR and XDR rates, HIV+/-
    - UNICEF Malaria data https://data.unicef.org/topic/child-health/malaria/ - ACT treatment used as treatment coverage for p. falc & p. vivax calculations

Clean/Double Checking of Data

- Maintain clear citations of data taken from databases. Citation should include author, source title, date, company, year, website, etc. A copy of the source should also be saved as a pdf and stored in the disease's references folder. In the citation, provide links to both the original website and the google drive document, in case the original link goes dead.
- Have at least one or two group members check any alterations, especially when shifting locations of cells/ranges; this can cause referencing errors that are hard to track later.
- when downloading data from a source, it may include more years than you are searching for. To make it easier to check data, apply a "filter" to the sheet; this will be one of the options

under “data” in excel and google sheets, and you can choose to show specific years, for example.

- It often is easiest to copy and paste the data you are checking into a separate excel or google doc sheet side by side with the database info, make sure to include countries to check if they match up too

Implementation Of Efficacy Data

The modeling team works hand in hand with the SR team to implement efficacy data into every disease model. Since the SR team is not as familiar with the format of the models, it is essential that the two teams communicate to ensure that the efficacy sheets are properly referenced in the modeling code so that the correct data is imported into the ORS. In order to make the process go smoother, the modeling team has created an explanation sheet that details what aspects of the efficacy sheets must remain the same from semester to semester in order to prevent the breaking of the modeling team’s code. A few properties to focus on is keeping the names of the efficacy sheets the same, or at least notifying the modeling team that the name is being changed so the import range code can be fixed, and putting columns that are referenced in the models in the specific location as listed on the explanation document.

To better understand how efficacy is implemented, please refer to this training assignment. Which goes through how to pull the efficacy data from the SR teams sheets and implement them into the models.

All efficacy data is transported from SR files to the modeling ORS’s through an IMPORTRANGE formula. This assures that any update made by SR will be transferred to the correct modeling documents. NEVER copy and paste SR sheets into the modeling efficacy sheets because the updates made by SR will not be mirrored in the modeling sheets/actual models.

Sensitivity Analysis

- Purpose: Create numerous copies of our existing Overall Rating Sheets. Manipulate some parameter or variable to assess how much influence that parameter has on our rankings. Do some companies shift spots if we weigh variables differently?
- Spreadsheet located in ORS 2015 or ORS 2020 folder. Lists assumptions made for each of the models and the locations of those variables.
- Copies of the models will be made and those assumptions manipulated to observe any large resulting changes on scores, rankings, etc.
- Instructions on how to do sensitivity analysis

Expert opinion / Literature Review

This team, while under the guidance of Professor Hassoun, is ultimately student-run. While anyone can make mistakes, students are more likely to make them, as they are not experts on the work we are conducting. Therefore, it is often necessary to consult more experienced professionals when necessary to ensure the work we are publishing is accurate and correct.

Getting an expert opinion is often a case-by-case basis task. Therefore, it will likely be performed under the guidance of Professor Hassoun; you should contact her throughout this process to know how to proceed.

That being said, in every case, it is necessary to follow proper etiquette. You are representing the organization when you talk to external professionals. Therefore, make sure to use appropriate language when communicating, and express gratification for their efforts.

Important Contacts

Sake de Vlas - s.devlas@erasmusmc.nl . Author of: Concerted Efforts to Control or Eliminate Neglected Tropical Diseases: How Much Health Will Be Gained?

Paper writing

Current examples

Alongside the information presented on our website. The GHI project has also published papers in journals presenting the information we found as a team. Some examples you should be familiar with are the PLOS paper and TMIH paper.

The PLOS Paper is the original paper written for the project, linked to the Old Website Model (V. 63 in ORS 2014). It has a slightly outdated Impact Formula, but the general format is standard for writing papers. It only covered HIV, TB, and p. falc malaria.

The TMIH paper is a more current report. It follows a similar format to the PLOS paper and goes through the TB, HIV, Malaria, and NTD models.

Other examples:

In 2022 we worked on completing the NTD paper and supplement. The Hep C paper is also in process as of winter 2023, and is anticipated to be published soon.

Another important task that the modeling team worked on this semester was to write a sources paper that denoted our methodology for source impact allocation, determining lives saved by a given source (company), daily dose information, as well as determining disease burden alleviation

by a given source. This may also be a helpful resource for an example of how to format papers in the future.

Advice on writing papers

The previous examples above serve as good examples of what new papers should look like. Papers should be general enough for anyone to follow along with a model and understand. Avoid in depth explanations of Excel formulas, instead focusing on the overall impact formulas/choice of variables.

.

Checking papers

The writing process can sometimes take a long time. In that time, errors can be found in the model that need to be fixed, a report posted to the website sometimes needs to be updated with the most current model, or a new set of eyes is needed to verify that all of the information was put into the paper correctly. In these cases, someone needs to go through and check each of the numbers of the paper and make sure they were all put in correctly. While this might seem like a straightforward process, there are some good practices to ensure that you check the information correctly.

1. **Use suggestion mode while making changes:** Google docs has a feature where instead of directly making edits, you can suggest them. This way, if you made a mistake while checking, it is not permanent, and the original value is clearly listed.
2. **Comment what sheet/tab/cell you got your value from:** When you suggest an edit, also make a comment or reply to the comment made when you suggest an edit with the information necessary to find where you found your value. Make sure to link the sheet you were looking at, the tab the value is in, and the cell, or the cells you added to find the value you were looking at. This way, when your team leader looks at your edits, they are easily able to see where the values you calculated came from, and verify if you are correct.
3. **Make a copy of the sheet for calculations:** Sometimes, the values in the paper are not exact values in the cells of the sheet. For example, the NTD paper refers to the total impact of NTD drugs in South America. The ORS does not have a cell that just sums the impact of South America, they have one for all of the countries, or each of the individual countries. To easily calculate this, you can make a copy of the sheet and add an extra tab just for calculations. Then, you can take the sum of just the countries you are looking at.
4. **Pay attention to wording:** Sometimes, the wording in an academic paper or report is not the same as the ORS. For example, "need" may not be called "need," but rather "DALY's accrued absent treatment," and the "impact score" might be called "DALYs alleviated." If the wording is ever unclear, make sure to reach out to your group leader to make sure you are looking at the correct information. It is better to send an email and clarify, then spend time looking at the wrong information and have to redo it.

Writing Model/Technical Guides

As a modeling team member, you should get familiar with the technical guides. They provide explanations on each of the columns in the models.

Key Powerpoint Elements:

- General Info with link to spreadsheet and description of the first basic columns (those without calculations that simply have exogenous information).
- Explanation of different strands of the disease (if applicable)
- Column by column explanation of respective variable and calculation of respective variable.
- What DALYs are/mean and how to calculate them (since our models just list them as given from a source)
- Examples of every calculation. Try to keep example consistent within the presentation, for example always use the same country across all numerical examples so that in the example you always reference the same cell
- Miscellaneous definitions of things such as who the WHO are, what efficacy is, etc.
- Note of what different abbreviations stand for
- Screenshots from excel with pointers towards which columns we are looking at and in what order we use them in in our syntaxes
- Maybe better to not copy and paste whole syntaxes in, especially if the syntaxes are a series of looking up different data such as regional if country is not available, and world if regional is not available. Instead of a long syntax, that is overwhelming to the reader and that the reader probably won't read carefully anyway or know where to start. We can explain the repetitive syntaxes in our own words.
- Include sources- example: WHO website
- More consistency in using column names

Suggested Powerpoint Elements:

- Editing all power points to have the same format (gray background, border, and fonts) as the ones currently on the GHI site.
    - Schist, Malaria, and HIV would need to change
- Graphs, for those models for which we have sufficient information

- How is our model distinct from other models (?)

Notes from Debbie for Power Points

-Explanation of sources of data, where do we get DALYs? - WHO, GBDx, etc.

-Definitions of applicable terms - DALYs, prevalence

Website Checking

The website checking process is an important step in publishing our results on our website. The modeling team needs to check every value on the website to ensure it matches the ORS, and work with the CS team to make the appropriate changes.

How to Check the Website

While it is not a difficult process, it is time consuming. Modeling team members have developed methods for streamlining the process. The standard procedure for checking the website is contained in this document. Make sure you have access to the most recent copies of the ORS to ensure you are checking the most recent numbers.

Reet made a video explaining how to check the website in the spring of 2022. While there have been changes to the process since the video was made, it helps demonstrate how to find the information.

Make sure to use this doc to learn how to check the whole website at a much faster pace.

Common Challenges

A challenge people face in the website checking process is not knowing where to find the information in the ORS. This can come from inexperience working with the models, or inconsistencies in naming. For example, 'Need' is called 'Disease Burden (DALYs) Absent Treatment' on the website. If you are ever confused or unsure, refer to the technical documents for further clarification, or reach out to your team leader. It is important to make sure you are checking the website against the correct numbers.

Standard format of long and short website check

Lastly, since this task involves working with the CS team, it is important to make sure the materials shared between the teams are clear and easy to read. It is important to use a standardized format for checking the website and make sure all members understand how to use it. Here is a long standardized template for checking the website. Note how it has spaces for the:

- Index (Country, Company, Drug, Disease, Source)

- Data type (impact score, need, etc.)
- Value from ORS
- Value from website
- difference
- ORS cell
- ORS tab

It is important to include all of this information to make sure the errors are appropriately reported to the CS team.

If you are chosen to do a quick check of the website, follow this standardized template. Once done with each section, make sure to change the color of each tab to green. Use this doc for help starting.

Creating New Models

Over the past few years the modeling team has attempted to create models for new diseases not originally covered in our ORS's. This includes trachoma, hepatitis C, and leprosy (which is estimated to be in the works during the summer or fall of 2021. It is often team leaders or very experienced members that are tasked with creating new models, but it is important to provide some general guidance. Here is an example of general steps we will take to complete the leprosy model. Treatment guideline research, efficacy data implementation, finding manufacturing data, writing tech docs and/or powerpoints, and possibly a writeup are all other important factors when creating a new model that can be replicated in the future.

Checking models

Whether you have created an entirely new model, or made an updated version of an old model, you may need to check a model at some point. Here are some ways to do this. The best method is a combination of both.

Recreating the model

When you first make a new model, you should have someone else recreate at least parts of the model to ensure you implemented the formulas correctly. Make sure they know what the basic ideas for your formula are, and have them recreate them without looking at your formulas.

The information you found, like the DALYs etc. can remain the same as someone recreates your model. They might not even have to recalculate everything, just the areas that are more prone to error, like efficacy and impact score.

Hand Calculate a few rows

Another method for checking a model is to hand calculate a few rows. This is something you can do yourself, as well as have someone else do as well. To do this, you want to pick a few rows and calculate the impact score for each. When you have finished, compare what is present in the model. If they match, the values must be calculated correctly.

Here is an example of doing so for the Hep C 2015 model in the ORS 2022(2015/2017). I hand calculated the impact score for SOF+DCV on G3 (acute) infections of Hepatitis C in Uganda.

Acute DALYs: 494.58

G3 distribution: 2.90% (regional value for AFR in cell Q229)

G3 acute DALYs: 494.58*0.029=14.34282

Treatment coverage: 2.20% (regional value for AFR found in cell D227)

Efficacy: 92% (WHO value Z227)

Impact score: (14.34282*0.92*0.022)/(1-(0.92*0.022))=0.2962957018

Needs to be: 0.2962927064 (Cell DC205) Correct! (rounding difference)

Automation Project

In an effort to streamline making models in the future, we have started automating them. This will reduce the amount of time spent creating the models. As the modeling team, our responsibility in this project is checking the automated models and making sure they are accurately pulling data from the sources.

Information on this process can be found in this folder. The contents are as follows: how to check an automated model and creating outlines. Checking the automated model is where we pull data directly from the source and compare the data in an automated model. This is done to make sure the correct values were pulled. The outlines are the shell the automated model fills in. For those who have never constructed a model themselves before, the first rows, serving as titles for each column, are usually the same for each model. Usually when you make a new one, you can copy and paste these columns in and fill the rest in accordingly. This is similar in the case of the automated models, except rather than filling them in yourself, they are automatically completed.

A link to the automation demo video by Alex is linked here:
https://doc-0k-4s-docs.googleusercontent.com/docs/securesc/2g6d30v51uengacfkgpndv3l10magibc

/3mduv0rmnh9ajnmfmf1ialiqbqoqrre3/1655678175000/15685296364465309682/15685296364465309682/1DL6zm7gLXaGDpnOm4SQP8gFM1Fte_fix?e=download&authuser=1

In order to check the automated model against the original source, you need to first know what the source is. Not all of our information comes from the GHDx website. To find this information, you can use the automated models sources sheet. This video goes in depth on how to use this sheet, but essentially, the columns all line up to a column in the ORS, and have the source from which the information comes from linked.

More information on the automation project can be found here, in Anna's own words.

Making Automated Models

Since making automated models is its own extensive project, including work from the CS, modeling, and SR teams, it was necessary to make a separate Automation Handbook. This handbook contains all of the information needed to use the automation tool to input automated data, as well as find non-automated information.

Forecasting Tool

The goal of the forecasting tool is for users to be able to "forecast" the impact score according to their specifications. They can use this to see what impact score would be generated based on the country, disease, and pharmaceutical company selected. While this is built by the CS team, it is the modeling team's responsibility to check that the data they are using are in accordance with the ORS data.

Instructions on how to check the forecasting tool data: Checking Forecasting Tool for Modelers

For more information on the forecasting tool, see the "Forecasting Tool Manual" and Forecasting Tool 2020 Folder

As of May 2021, the newest version of the forecasting tool manual has been updated to reflect the new forecasting visualizations that will be displayed on the website. These updated visuals were made possible by the new 1:1 country to drug table, which makes the impact scores specific to company, country, and disease, and allows the CS team to implement specifications readily on the website. The tabs labeled "For Modeling" are the result of an import range formula from our ORS and may need to be checked occasionally to make sure that the modeling data is being correctly reflected on the website. The static numbers are meant specifically for CS.

EALY and DALY Models

In Spring 2021, the modeling team worked on generating versions of the ORS that utilized EALYs rather than DALYs. All of the information (models and accompanying documents) can be found

in the 'EALY ORS' folder. The CS team has been given the documents in the folder titled "Compilation Data For CS".

The EALY model was made before the Summer and Fall 2022 edits, and therefore do not reflect the most recent versions of the ORS

What are DALYs?

To understand EALYs, we first have to understand DALYs. DALYs stands for "Disability-Adjusted Life Years." It essentially takes into account not only the death toll of diseases, but also the years impacted by disability. This provides a more complete picture on the impact of a disease in the area, as death is not the only result of disease.

Some resources that help further illustrate what a DALY is is "The DALY Show, Disability-Adjusted Life Year (DALY)", and "Burden of Disease: What does DALY mean?, as well as this page on the WHO website. We also have a training assignment, "How to calculate DALYs," which also explains how to do so.

The problem with this measure of impact is that it equates all years lost to disability. Loss of ability can exist on a spectrum, and diseases may result in different losses of ability. Therefore, EALYs were created to account for this difference.

What are EALYs?

EALYs are another method of measuring disease burden. The reason for using EALY figures over DALY figures is to be able to more equitably value the lives of those that are healthy with those that are disabled.

Mathematically the difference between the EALY and DALY model is that DALY's are years of life lost (to premature mortality)(YLL) + years of life lost to disability(YLD), EALY's are YLL*x + YLD, where x is any integer > 0 in order to account for different weightings of the YLL .

In Spring 2021 were generated the EALY folder where we replace EALYs with DALYs. On top of that, we change how EALYs are calculated to be EALYs = 2*YLLs + YLDs, EALYs = 3*YLLs + YLDs... EALYs = 10*YLLs + YLDs. These results intend to be implemented as an option on the website.

Making New Training Assignments

As a member of the team, you may be asked to make a new training assignment. This is usually done by the team leaders, but if you have become specifically proficient in a certain area, you may be asked to create instructions on how to recreate your work.

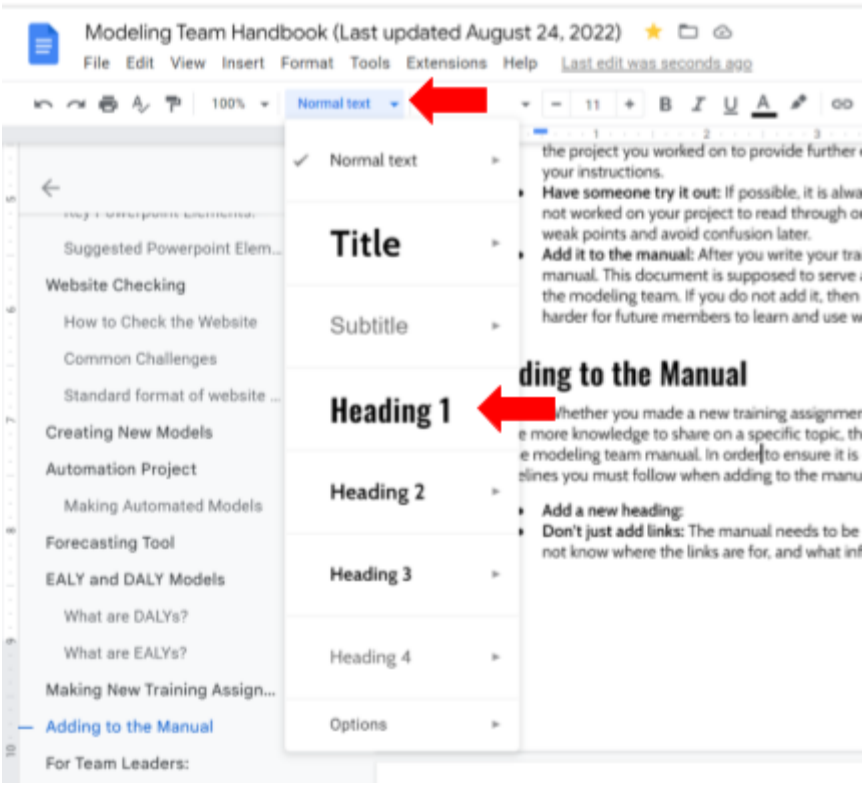
This manual is filled with examples on different training assignments that have been created throughout the years. You can use these as blueprints when writing your own instructions.

- **Explain it like you would to a middle school student:** You cannot assume that the person reading your instructions would be able to understand it at the level you do. Maybe they are unfamiliar with excel, or using the websites you use regularly. Write your instructions as if you were giving them to a middle schooler so there is little to no confusion.
- **Add links for further resources:** Sometimes, concepts are explained better elsewhere, or getting into detail may be unnecessary for the majority of your readers, and take away from your instructions. In these cases, add any links you have found helpful in your development of the project you worked on to provide further explanations to the future members reading your instructions.
- **Add pictures and screenshots:** If possible, add pictures to explain what you mean. It can add further clarification, and make the section look less daunting.
- **Consider making a video:** If you require a lot of explanation in your instructions, consider making a video. This can sometimes make it easier to understand, as you can demonstrate where things are located on your screen as you explain it.
- **Have someone try it out:** If possible, it is always good to get another member who may have not worked on your project to read through or try out your instructions. This can help identify weak points and avoid confusion later.
- **Add it to the manual:** After you write your training assignment, it is important to add it to the manual. This document is supposed to serve as a place to find any information you need on the modeling team. If you do not add it, then it might get lost or forgotten. This will make it harder for future members to learn and use what you have made.

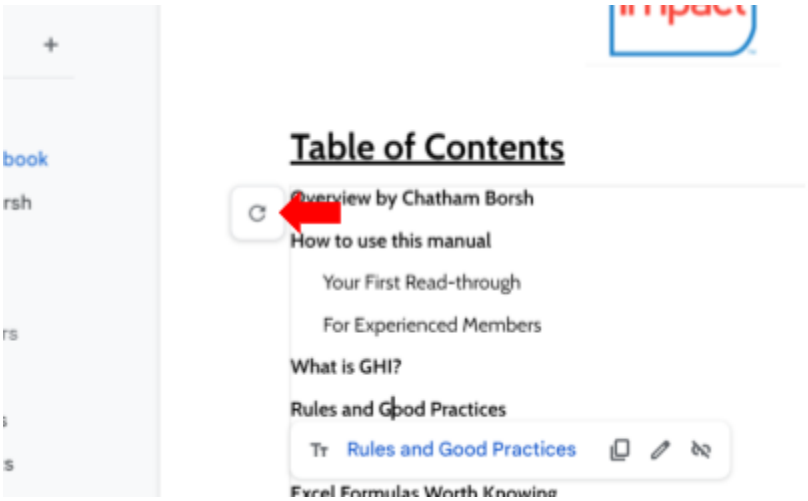
Adding to the Manual

Whether you made a new training assignment, have found a forgotten old one, or just have some more knowledge to share on a specific topic, there will come a time when you may need to add to the modeling team manual. In order to ensure it is easy to read in the future, there are some guidelines you must follow when adding to the manual.

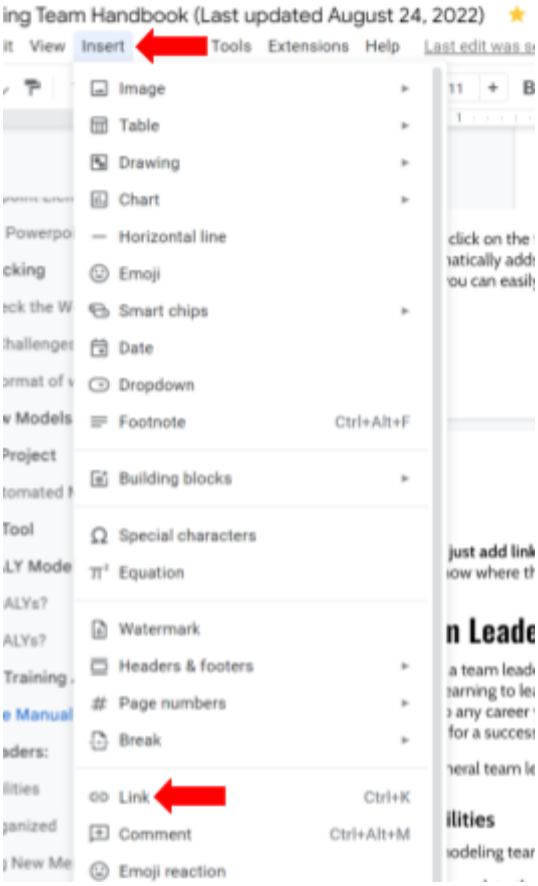
- **Add a new heading:** When you add a new section, don't just change the font size to match the headings already present, actually make it a heading 1, 2, 3 etc.



This doesn't just make it visually match the rest of the manual, but also makes it easier to add to the table of contents. If you click on the table of contents, and then the icon next to the top left corner (as shown below), it automatically adds the new heading and updates the page number where it is located. This way, you can easily add your new section anywhere in the manual.



- **Don't just add links:** The manual needs to be readable. If you just add links, the reader may not know where the links are for, and what information they contain. Make sure to introduce any assignments or resources you add with a quick explanation of what you will find. This way, people do not have to click through a dozen links to see if it contains the information they need. You can also go to insert > Link to add a link to a word in the document, making it look clean.



- **Consider where you are adding information:** Just tacking sections on to the end of the manual might not always make the most sense. Try to consider where things are located in the manual, and add to more similar sections. This way, future members can find important information in one location, rather than multiple. It also reduces the risk of your addition being overlooked or forgotten. You may also want to add to an existing section, rather than making a completely new section. In This case, you can use a Heading 2 or Heading 3 to add a subsection to an existing, large one.

For Team Leaders:

Being a team leader takes a lot of work. It's a lot of responsibility but also a ton of great experience. Learning to lead and work with a team of people for a common goal is a skill that will

transcend into any career you end up pursuing. So great job being in the position you are in and below are some tips for a successful semester:

For general team leader guidelines do be sure to check out the Group Leader Handbook

Responsibilities

The modeling team leader has a two key responsibilities:

1. **Assign work to the modeling team:** The modeling team leader is responsible for ensuring that each team member has 10 hours of work each week. Therefore, it is your job to plan out what needs to be done that week and delegate tasks accordingly. Make sure to reach out to Nicole to understand her expectations for the team, and prioritize her requests for the team.
2. **Train new members:** The modeling team leader also is tasked with training any new members who join the team. This involves designing a training plan as well as being available to answer questions when needed. You are the first person your members contact when they need help with something, and it is your responsibility to to

This section of the handbook was designed to help you navigate how to do this, and stay organized in the process.

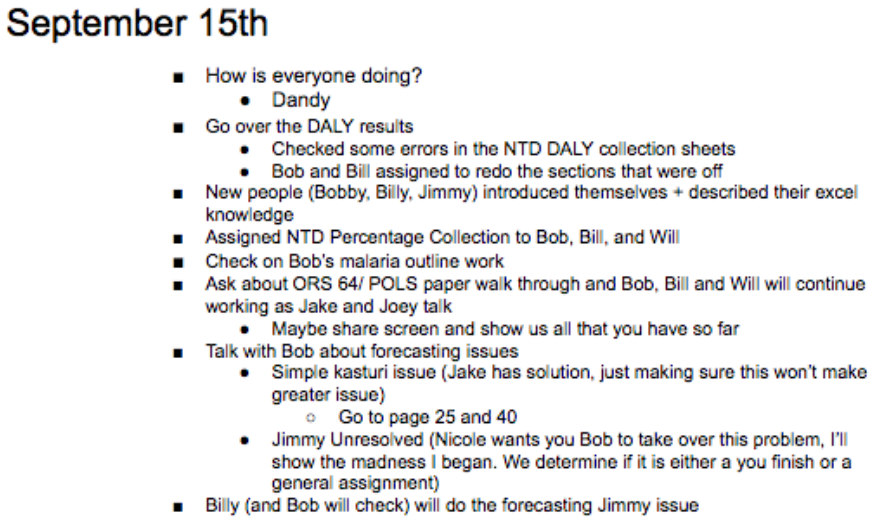
Staying Organized

As modeling team leader, you take on a lot more responsibility than you previously had. You are now not only responsible for your own work, but the entire modeling team's. Here are some tips on how to stay organized:

- **Establish overall goals**: Before each semester, you should identify current progress and previous semester achievements as well as create a plan for the semester. This is important for you to establish with Nicole before the semester begins. Having goals to work toward will help you stay on track and plan out everyone's assignments each week.

- **Keep meeting minutes:** A lot of information is shared during a GHI meeting. This is your time each week to communicate directly with Nicole, as well as your entire team at once. Therefore, it is necessary to keep track of the information shared during this time to make sure nothing is lost. A meeting minutes document can be prepared ahead of time as the agenda for both the class and team meetings. It is also recommended that you use this document to assign tasks to each member for the coming week. This way they have something to refer to in case they don't keep track of it themselves. It is recommended that each meeting gets its own page. As the meeting progresses, assign someone to keep minutes on that meeting's page.

  Here's an idea of what your minutes doc should look like:



September 15th

- How is everyone doing?
  - Dandy
- Go over the DALY results
  - Checked some errors in the NTD DALY collection sheets
  - Bob and Bill assigned to redo the sections that were off
- New people (Bobby, Billy, Jimmy) introduced themselves + described their excel knowledge
- Assigned NTD Percentage Collection to Bob, Bill, and Will
- Check on Bob's malaria outline work
- Ask about ORS 64/ POLS paper walk through and Bob, Bill and Will will continue working as Jake and Joey talk
  - Maybe share screen and show us all that you have so far
- Talk with Bob about forecasting issues
  - Simple kasturi issue (Jake has solution, just making sure this won't make greater issue)
    - Go to page 25 and 40
  - Jimmy Unresolved (Nicole wants you Bob to take over this problem, I'll show the madness I began. We determine if it is either a you finish or a general assignment)
- Billy (and Bob will check) will do the forecasting Jimmy issue

- **Maintain folder security:** Make sure you clearly document the location of, and transfer ownership of each document to Nicole. Your modeling team members do not have access to the rating systems folder, so it is your responsibility to make sure everything is in a clearly labeled location.
  - **Practice folder security with your team:** While your team members do not have access to the rating system folder, you can still train them on how to maintain organization. One option is to make a folder for each new member in the training folder. This way, they can practice putting their work in this folder and transferring ownership. You can even make a smaller version of the Folder Tracking Document to have them practice noting where the information is.
- **Create a priority list:** When you have so much on your plate, it is important to establish a priority list to make sure nothing falls through the cracks. Write down any assignment you are given by Nicole, and put them in order from most pressing to least. Sometimes, priorities shift throughout the semester, so it is recommended to make a list in google docs and keep adding to it and moving things around when needed.

Introducing New Members to Modeling

Introducing new members to modeling and training them is the most important task you take on as a modeling team leader. Creating a healthy, knowledgeable team will ensure the future of the project, and it is your job to foster this.

- **Be understanding:** A lot of people join the team with little to no experience with modeling or excel. It can be overwhelming to start. Make sure your team knows that they are not under-qualified to learn how the models work or participate in the modeling team. Everything looks harder than it actually is, do your best to support them through the early learning curve.
- **Reach out early and often:** Just like they have to reach out to you, you need to reach out to them often. Check in daily on their progress and offer a chance to ask questions. A lot of

times, people will avoid reaching out, but if prompted, actually have a lot of questions. Reaching out first can make them more comfortable coming to you when they need help.

- **Set up separate meetings:** Meetings with Nicole and the whole team can get confusing for new members. Nicole often has a lot of people to talk to, and you may be hearing information for the first time as well. Setting up a separate meeting, usually over zoom, during a time everyone can meet (using tools like when2meet is good for this) can create a calmer environment where you can more easily communicate with your team and explain tasks.

The first few weeks:

You should plan out the first few weeks of assignments for all of your team members. One method that has worked in the past is establishing a training plan. The Spring 2021 training plan and Fall 2022 training plan are both good examples of what training plans can look like. You can design this however works for you, but it is good to include some key elements:

1. **Reading the manual:** As of winter 2022, the modeling team manual has been designed to share all of the necessary information the new members need to learn to be successful on the team. They should read it within the first week of being on the team. You can either assign the whole manual, assign the handbook training assignment, or assign sections of the manual to go through that you believe are most important. There are training assignments throughout the manual as well, so make sure your members are aware of this, and complete them in the first few weeks.
2. **Excel formula practice:** It is hard to understand the models without first understanding the formulas, therefore, it is important to assign this task early in their training, so they have some knowledge to reference when looking at the ORS.
3. **Reading the PLOS and TMIH papers:** The PLOS and TMIH paper training assignments can serve as good introductions to where information is located in the models. When assigning these tasks, make sure to demonstrate how to read the papers so your team members can get comfortable referring to them.
4. **Introduce the models and formulas in person:** Taking the time to explain all of the models and formulas over the course of your first few meetings can help clear up any confusion with your team. Here is a good example of the order you can do this in:
    a. Go through the Company, Impact Score, DALY, and Need tab for the first week. These are simple enough tabs and should be explained from the beginning.
    b. Mention the conceptual model, with the boxes. Don't go too in depth on this at first, as it tends to be confusing for someone new to the project. Note how it then connects to our Impact formula, I = DeΘ/(1-eΘ), explaining each variable. The resulting value will have the same units as DALYs.
    c. Bring up one of the simpler models, such as Onch or Malaria. Point out key details and columns, including DALYs, treatment coverage, and efficacy. Note how we fall back on regional or global averages when lacking country-level data.
    d. Expand a few of the formulas, explain some syntax if they're less familiar with Excel. Especially open an impact formula, point to our section that is our impact formula, ABC/(1-BC).
    e. All models except TB follow the same general left-to-right layout. Compare across models. Give them one to work across, checking what columns reference each other.

5. **Model Reconstruction:** While you can explain the models all you want, your team will never fully understand them until they work with them. Model reconstruction is the best way to do this. Instructions have been created on how to reconstruct the Malaria model, but you can also ask members to reconstruct other models using the tech documents. Since this task is often used for teaching how to implement the formulas, it is usually best to have them recreate an existing model, as the DALY and Efficacy information has already been found. Finding new information is often gratuitously time consuming, and takes away from members learning how the models work.
6. **Use existing training assignments:** The ORS 2020 folder has a plethora of training assignments that have been created over the years. Take a look at your goals for the semester, and assign the training assignments you believe would be most applicable to your specific needs as a team. Make sure to explain each assignment thoroughly before assigning it to your team so they feel comfortable completing it. (try to also assign any present in the handbook by its page in the handbook, as it also comes with other supporting information).

Moving Forward

- **Be ready to answer any questions:** Make sure to understand your task before you assign them. Especially in the beginning, you need to make sure that you can explain the tasks you assign thoroughly and give them a clear idea of what to get and where to put it.

- **Trust your team:** When you get a more confident team, trust them with harder tasks. Remember, you are training the next team leader, they need to be as competent as you are. Delegate as many tasks as possible to your other team members, and trust them to learn how to contribute to the team.

- **Don't shy away from hard conversations:** You need to trust your team, but your team also has to be trustworthy. If they are not completing their hours, or doing the work you ask of them at the quality you require, it ultimately reflects back on you. It is your job to have a conversation with them and reestablish the expectations, and terminate members if needed. It wastes your time if you have to redo their work, or if you cannot trust them to take on tasks. Always give people the benefit of the doubt, but also know when to be firm.

- **Refer back to older semesters:** If your semester calls for creating more up to date disease models, please use this resource: https://drive.google.com/drive/u/0/folders/1g9wOaNFqXM_j-P1yN8gyyKrCByywUNAu. The 2020 Fall Model team spent a majority of their semester writing detailed instructions on how to update these disease models... no need to reinvent the wheel. Make sure to give that all a look and consider it with future years.

- **Demonstrate appreciation for their work:** It is crucial to make sure that the modeling members know what they are doing and how it matters. The modeling team can feel a little lost in what they are doing and whether it matters. Remind them that it does. Show them the impact of their work and the organization as a whole. Looking at the website can be a good simple way to show them the results of their work.

- **Reach out if you need help:** Feel free to email Jake West at westjacob426@gmail.com if you are having trouble with establishing your semester plan and execution. Take a look at our planning ahead folder to see how previous leaders have guided the transition of power between semesters. And as always, reach out to Nicole early and often for support.