Genomics costing tool

User manual

APHL Association of Public Health Laboratories

FIND Diagnosis for all

The Global Fund

UK Health Security Agency

World Health Organization

Genomics costing tool

User manual

World Health Organization

Genomics costing tool: user manual

ISBN 978-92-4-009086-6 (electronic version)
ISBN 978-92-4-009087-3 (print version)

Design and layout by Antonio Perez.

Contents

Acknowledgments

The Genomics Costing Tool (GCT) alongside its user manual was codeveloped by five institutions: the Association of Public Health Laboratories (APHL), the Foundation for Innovative New Diagnostics (FIND), the Global Fund to Fight AIDS, Tuberculosis and Malaria (TGF), the UK Health Security Agency (UKHSA) and World Health Organization (WHO).

Leadership and coordination

Leadership and coordination for the development and publication of the GCT was provided by WHO, with coordination and technical lead from Joanna Salvi Le Garrec (Zwetyenga) (Infectious Hazard Management, WHO Regional Office for Europe, Copenhagen, Denmark) and Oluwatosin Wuraola Akande (Department of Epidemic and Pandemic Preparedness and Prevention, WHO, Geneva, Switzerland).

Contributors

The following members of the GCT working group made substantial intellectual input by contributing to the technical concept, development, writing and review of the GCT and its user manual: Babak Afrough (UKHSA, London, United Kingdom of Great Britain and Northern Ireland), Oluwatosin Wuraola Akande (Department of Epidemic and Pandemic Preparedness and Prevention, WHO, Geneva, Switzerland), Maria Valerie Amante (Infectious Hazard Management, WHO Regional Office for Europe, Copenhagen, Denmark), Lisa Carter (Country Readiness Strengthening, WHO Lyon Office, Lyon, France), Noah Hull (Association of Public Health Laboratories, United States of America), Alexandr Jaguparov (Infectious Hazard Management, WHO Regional Office for Europe, Copenhagen, Denmark), Marco Marklewitz (FIND, Geneva, Switzerland), Biran Musul (WHO Türkiye Country Office, Ankara, Türkiye), Ashley Norberg (Association of Public Health Laboratories, USA), Angela Poates (Association of Public Health Laboratories, USA), Swapna Uplekar (FIND, Geneva, Switzerland), Toni Whistler (TGF, Geneva, Switzerland), Aude Wilhelm (UKHSA, London, United Kingdom), Joanna Salvi Le Garrec Zwetyenga (Infectious Hazard Management, WHO Regional Office for Europe, Copenhagen, Denmark).

The following provided support to the scoping and review of the tool: Céline Barnadas (Country Readiness Strengthening, WHO Lyon Office, Lyon, France), Sebastien Bruno François Cognat (Country Readiness Strengthening, WHO Lyon Office, Lyon, France), Jane Cunningham (Department of Epidemic and Pandemic Preparedness and Prevention, WHO, Geneva, Switzerland), Leena Inamdar (UKHSA, London, United Kingdom), Anila Kreku (UKHSA, London, United Kingdom), Dmitriy Pereyaslov (Department of Epidemic and Pandemic Preparedness and Prevention, WHO, Geneva, Switzerland), Gina Samaan (Department of Epidemic and Pandemic Preparedness and Prevention, WHO, Geneva, Switzerland), Anita Suresh (FIND, Geneva, Switzerland).

External reviewers

The GCT working group would like to thank the collaborative efforts of all the public health laboratories that participated in the pilot exercise that contributed to validation of the GCT.

African Region

Rexford Adade (National Public Health and Reference Laboratory Ghana Health Service, Accra, Ghana), Bright Adu (Noguchi Memorial Institute for Medical Research, Accra, Ghana), Ivy Sarfo Agyemang (Noguchi Memorial Institute for Medical Research, Accra, Ghana), Miriam Appiagyei (Noguchi Memorial Institute for Medical Research, Accra, Ghana), Kwame Asante (Association of Public Health Laboratories, Accra, Ghana), Gifty Boateng (National Public Health and Reference Laboratory, Ghana Health Service, Accra, Ghana), Kwame Boamah Buabeng (Noguchi Memorial Institute for Medical Research, Accra, Ghana), Guracha Guyo (WHO Country Office Ghana, Accra, Ghana), Francis Kasolo (WHO Country Office Ghana, Accra, Ghana), Nicksy Gumede-Moeletsi (WHO Regional Office for Africa, Brazzaville, Congo), Quaneeta Mohktar (Noguchi Memorial Institute for Medical Research, Accra, Ghana), Christopher Tamal (WHO Country Office Ghana, Accra, Ghana), Peace Uche (Noguchi Memorial Institute for Medical Research, Accra, Ghana).

Eastern Mediterranean Region

Hanan Al Kindi (Central Public Health Laboratories, Ministry of Health, Muscat, Oman), Yathrib Al Zakwani (Central Public Health Laboratories, Ministry of Health, Muscat, Oman), Amal Barakat (WHO Regional Office for the Eastern Mediterranean, Cairo, Egypt), Rajesh Kumar (Association of Public Health Laboratories, Muscat, Oman), Luke William Meredith (WHO Regional Office for the Eastern Mediterranean, Cairo, Egypt).

European Region

Altyn Iskakova (National TB Reference Laboratory, Bishkek, Kyrgyzstan), Gulmira Kalmambetova (National TB Reference Laboratory, Bishkek, Kyrgyzstan), Kaliya Kasymbekova (WHO Country Office Kyrgyzstan, Bishkek, Kyrgyzstan), Meerbubu Sydykova (National TB Reference Laboratory, Bishkek, Kyrgyzstan).

Abbreviations

APHL	Association of Public Health Laboratories
BSC	biosafety cabinet
CapEX	capital expenditure
FIND	Foundation for Innovative New Diagnostics
GCT	Genomics Costing Tool
GISAID	Global Initiative on Sharing All Influenza Data
L&W	annual licence and warranty fee (ONT)
MF	manufacturer (Illumina or ONT)
NGS	next-generation sequencing
ONT	Oxford Nanopore Technologies
PCR	polymerase chain reaction
QMS	quality management system
rxn(s)	reaction(s)
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
TGF	the Global Fund to Fight AIDS, Tuberculosis and Malaria
UKHSA	UK Health Security Agency
WHO	World Health Organization

Glossary

Amortization	The cost of equipment spread over the lifetime of the equipment. Important to consider when costing to ensure equipment can be replaced when the lifetime has expired.
BaseSpace	An Illumina cloud-based application for data management, storage and analysis.
Capital expenditure (CapEx)	Money spent on acquiring (and maintaining) fixed assets such as equipment.
Establishment cost	The total cost to start a new next-generation sequencing (NGS) laboratory, including all associated costs.
Global Initiative on Sharing All Influenza Data (GISAID)	Online public database used to publish and observe global NGS data, including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and influenza.
Incidentals	Minor expenses occurring in every sequencing run.
Licence and warranty (L&W)	An annual fee charged to receive updates for Oxford Nanopore Technologies (ONT) instruments.
Quality management system (QMS)	A system for documenting processes, procedures and responsibilities in place to ensure quality results.
Sample retest	Based on the laboratory sequencing success rate, the number of tests that need to be re-sequenced to reach the annual throughput selected by the tool user.
Sequencing instrument ("instrument" or "sequencer")	The specific instrument model used for sequencing (i.e., Illumina MiSeq, ONT MinION).
Sequencing platform ("platform")	The type of sequencing instrument, defined by manufacturer type (i.e., Illumina or ONT) and can include any instrument model (i.e., MiSeq, iSeq, MinION).
Terra.bio	A cloud-based tool for data storage and analysis using developed open-source pipelines with a graphical user interface.
Throughput	This references the amount of DNA or RNA that can be sequenced at the same time. Low or high throughput references the amount of data produced, scalability and speed.

Introduction

Building on genomic sequencing gains, the global momentum to further its application and the political impetus provided by World Health Assembly Resolution 74.7, WHO launched the Global Genomic Surveillance Strategy for Pathogens with Pandemic and Epidemic Potential 2022–2032 *(1)*. This Global Strategy provides a high-level unifying framework to strengthen genomic surveillance capacities so that there is quality, timely and appropriate public health action within local to global surveillance systems. In alignment with Objective 1 of the strategy (improve access to tools for better geographic representation), it is crucial to have a comprehensive overview of expenditures and actual costs of setting up, scaling up or running a genomics laboratory.

This Genomics Costing Tool (GCT) builds on the Laboratory Test Costing Tool of the Better Labs for Better Health initiative *(2)*. WHO developed the GCT in collaboration with the APHL, FIND, TGF and the UKHSA.

Purpose

The tool currently facilitates budgeting and resource mobilization for infrastructure, workforce, biosafety and quality assurance associated with SARS-CoV-2 genomic surveillance. Subsequent versions of the tool will be pathogen agnostic, so that it can be used to cost genomic surveillance more holistically.

Target audience

This tool will be useful to country, regional and global policymakers, health administrators and economists, laboratory directors, quality managers, donor institutions and other stakeholders engaged in genomic surveillance for priority pathogens.

Software requirements

The GCT is a Microsoft Excel-based tool developed by Microsoft for Windows, macOS, Android, iOS and iPadOS. The tool is compatible with Microsoft Excel 2016 or newer. Users must have the following software installed: Microsoft Excel 2016, 2019 or Microsoft 365. The file is compatible with the following operating systems: Windows 10 or later and MacOS 10.12 Sierra or later. Internet connection is not needed to use the tool; however, it may be required for downloading Microsoft Excel updates or external data sources (if applicable).

These system requirements are provided to ensure a seamless experience when using the tool. It is essential to meet these requirements to fully use the tool's features and functionality.

Additional notes

Users should be familiar with basic Microsoft Excel functions and formulas. To protect the tool from potential threats, ensure that the system has up-to-date antivirus software. To prevent data loss, regularly back up the Excel file.

How to use the tool

General guidance on the tool

Input requirements

Items and costs requested from the user to use the GCT:

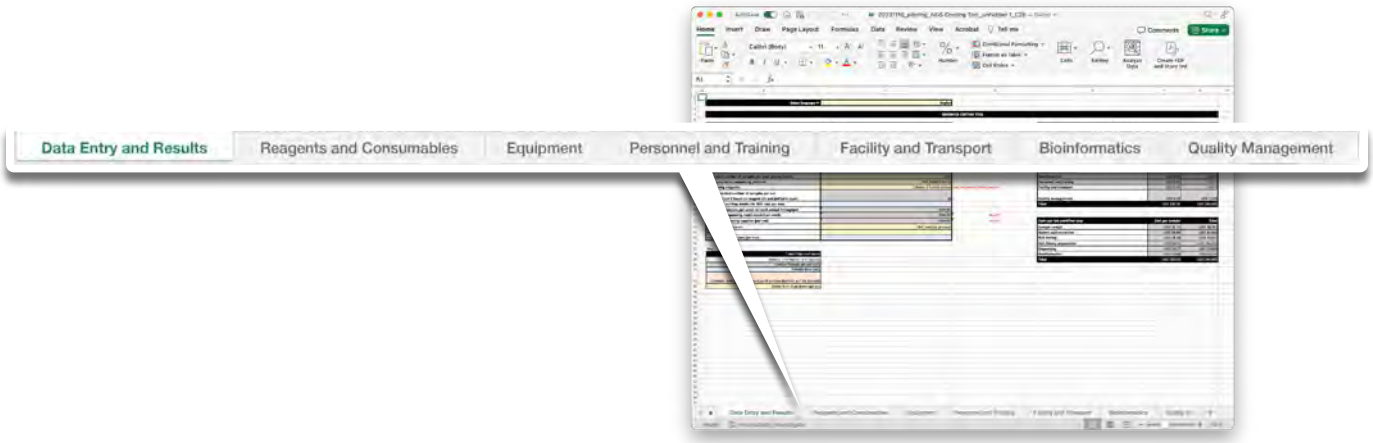
1. Number of SARS-CoV-2 samples sequenced in the last year:
   - If the last year's throughput does not accurately reflect the increase or decrease in recent throughput, use an annual estimate based on recent monthly throughput.
2. Sequencing platform(s) used for sequencing SARS-CoV-2.
3. Number of working weeks per year.
4. List of reagents and consumables used for SARS-CoV-2 sequencing workflows with prices (from sampling to analysing data) and quantities used in the last year.
5. List of all equipment used for sequencing SARS-CoV-2 (from sampling to analysing data), with equipment ages, prices and proportion of time the equipment is used specifically for SARS-CoV-2 sequencing activities (listed as a percentage).
6. List of all equipment maintenance contracts (from sampling to analysing data), with prices.
7. List of all personnel involved in SARS-CoV-2 sequencing and bioinformatics, with salary and training costs in the last year (from sampling to analysing data, if sampling is performed by a laboratory).
8. List of all facility-related costs with prices for the last year (rent, building maintenance, gas and heating, water, electricity, internet, telephone, waste management, generator maintenance, ventilation system maintenance, generator fuel).
9. Costs for sample transport related to SARS-CoV-2 sequencing in the last year, including any customs-related fees.
10. Costs for QMS (accreditation, external quality assessments, etc.) for SARS-CoV-2 sequencing purposes in the last year.

Some of the required information may need to be requested from the accounting, procurement or other supporting office.

Structure of the tool

The GCT is a Microsoft Excel spreadsheet with seven worksheets used to calculate the cost of sequencing SARS-CoV-2 based on annual sample throughput. A description of each worksheet is included in the manual, along with detailed instructions on the data entry requirements for each table in the respective worksheet (Fig. 1).

Fig. 1. Worksheets in the GCT



Note: Altering or renaming worksheets and cell labels will result in calculation errors and will compromise the results. It is not recommended.

Colour-coding

The tool is colour-coded for easier navigation. Table 1 describes the colours used with explanations. Notably, the grey-coloured cells are protected and cannot be edited as they contain formulas or fixed information. The peach-coloured cells are prefilled and some contain formulas; however, if a peach-coloured cell contains a different value from the user data, it should be changed accordingly.

Table 1: Colour-coding, content and user requirements of the GCT data cells

Color-coding, content, and user requirements of the GCT data cells

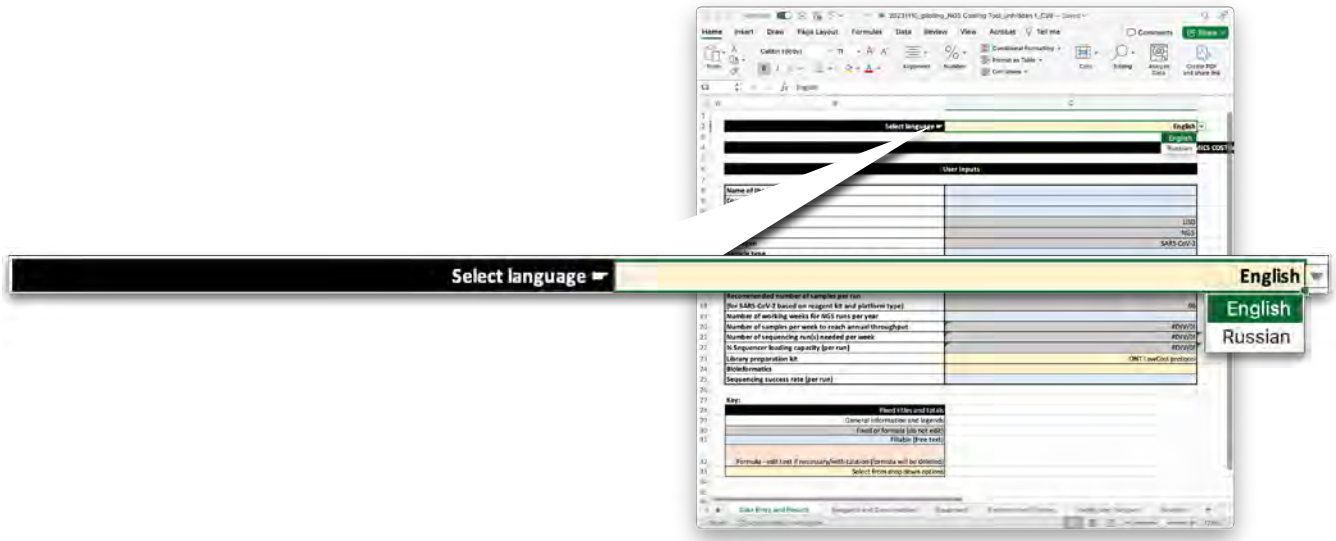
Colour	Description of the color-coding
Black	Fixed titles and totals
White	General information and legends
Grey	Automatically generated calculations/fixed information – protected cell that cannot be edited
Blue	Data to be entered by the user
Peach	Prefilled data. Data should be checked by the user. If information or values differ from what the user has, the cells must be edited to the correct information/values
Yellow	Drop-down menu available

Note: Prefilled peach-coloured cells may contain formulas. If needed, you may edit the content of the cell; however, doing so will cause the formulas to be lost, and you will need to use a new copy of the tool for costing a different scenario.

Language selection

The "Data Entry and Results" worksheet contains a drop-down menu for language selection. For the current version of the tool, English and Russian languages are available (Fig. 2).

Fig. 2. Language selection drop-down menu displaying English and Russian as available language



Additional considerations on the use of the tool

If the tool is being used to cost different laboratory scenarios (e.g., 600 vs 1200 sample throughputs or using one platform vs another), it is recommended to begin with a new, unused copy of the tool. If any formulas were erased during the initial use of the tool (i.e., to update actual quantities vs what the tool was auto-populating for any reagents), failing to use a new copy of the tool may cause errors in costing accuracy if those cells are not updated during subsequent use. Four examples of scenarios in which a laboratory may choose to use the GCT can be found in Annex 1.

Note: The drop-down menu selections in the "Data Entry and Results" worksheet do not refresh if a new platform or throughput is selected; thus, if the same file is used for a different scenario, the options will remain filled and must be manually reselected. Not doing so may lead to incompatible combinations (e.g., a reagent incompatible with a platform). If a new costing scenario is being generated in the tool, every line within the tool must be reviewed to ensure accurate costing.

This tool aids users and decision-makers in understanding the financial implications of implementation or continuation of NGS programmes . It is highly recommended that the tool be updated after any significant changes within the laboratory or to workflows (e.g., a bioinformatician is hired, salary increases are given to staff, new equipment is procured, the cost of reagents has changed). This will ensure that the tool provides the most reliable estimates.

Only one sequencing platform (i.e., Illumina MiSeq) can be accounted for at a time. If a laboratory uses more than one sequencing platform (i.e., Illumina MiSeq and ONT MinION) for routine sequencing of SARS-CoV-2, the tool will need to be filled out individually for each sequencing platform. One version of the tool would be completed in its entirety for the first platform. For additional sequencing platforms, open an unused copy of the tool and complete the "Data Entry and Results", "Reagents and Consumables" and "Equipment" worksheets, ensuring all data entered are specific to that platform. In particular, the throughput used for each individual platform should be noted in the "User inputs" table in each use of the tool. For the subsequent platforms, the "Personnel and Training", "Facility and Transport", "Bioinformatics" and "Quality Management" worksheets should be omitted, as these data will have been captured on the first completion of the GCT and that information will not change based on the platform used. For example, polymerase chain reaction (PCR) kits, biosafety cabinets (BSCs) and so forth that are used for the sequencing workflow would not need to be accounted for in the second version of the tool for the alternative platform and can be listed as quantity "0". To calculate the final cost for the laboratory's sequencing program me, compile the results from all the versions of the tool. In the future, an upgraded version of this tool will allow for multiple platforms within the same iteration of the tool. If a laboratory has multiple instruments of the exact same sequencing platform (i.e., two or more Illumina MiSeqs), this can be represented in the "Equipment" worksheet. Note that the message "platform with higher throughput may be needed" on the "Data Entry and Results" worksheet will not disappear if additional sequencing instruments are added to the "Equipment" worksheet.

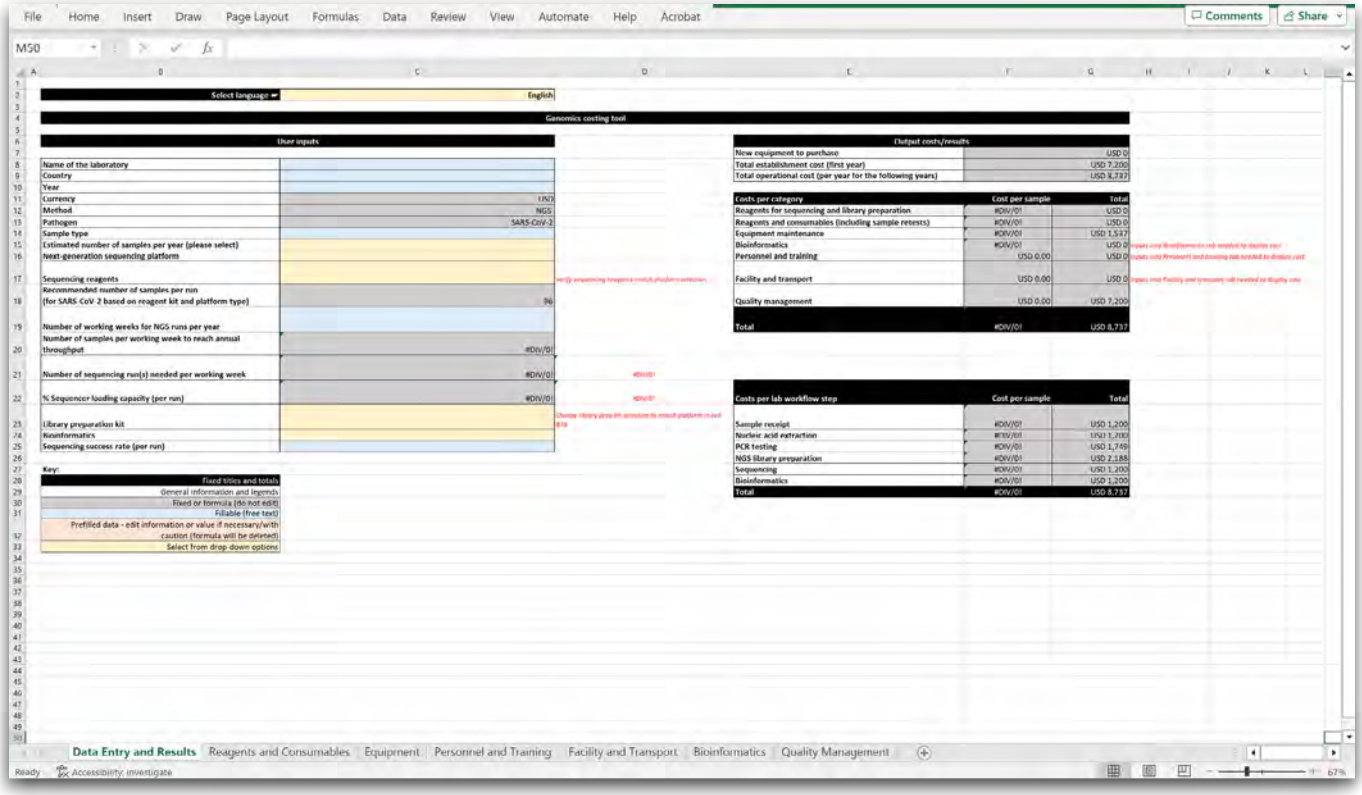
Inclusion of Illumina and ONT platforms

Illumina and ONT sequencing platforms are available in the GCT. The GCT working group surveyed data from GISAID to identify the sequencing platforms used. These are the two most frequently used manufacturers globally and were the most used platforms for the SARS-CoV-2 response.

Completing worksheets

Data entry and results

Fig. 3. Screenshot of the "Data Entry and Results" worksheet as it appears on opening the GCT



Note: Due to embedded formulas, some cells in the table will display "#DIV/0!" until the information needed by the tool has been input.

The "Data Entry and Results" worksheet contains two tables: "User inputs" and "Output costs/results" (Fig. 3).

Complete the "User inputs" table in this worksheet first. It captures basic information about the sequencing laboratory, which is used to pre-populate other worksheets and costs (Table 2).

The "Output costs/results" table displays final costs once all worksheets are completed. The results are broken into several sections for consideration (i.e., costs per category or by workflow step).

Table 2: Detailed explanation of each cell of the "User inputs" table in the "Data Entry and Results" worksheet

Worksheet: Data Entry and Results **Table: User Inputs**		
Data name	***User input requirement***	***Explanation***
Name of the laboratory	Enter the name of your laboratory.	
Country	Enter your country name.	
Year	Enter the year for which the costing is done.	
Currency	NONE	The currency used in the tool is USD.
Method	NONE	The method used in the tool is NGS.
Pathogen	NONE	The pathogen used in the tool is SARS-CoV-2
Sample type	Enter the sample type(s) which are used for testing.	E.g., nasopharyngeal swab.
Estimated number of samples per year (please select)	Select the throughput from the available options (600, 1200, 2500, 5000, 8000, 12 000) which fit the desired costing scenario – e.g., the number of samples sequenced per year or the estimated number of samples that you plan to sequence/scale up to.	Used to calculate total annual cost and cost per sample, as well as to auto-determine the quantity of sequencing kits, reagents and consumables.
Next-generation sequencing platform	Select the NGS platform from the available options. Depending on the throughput selected above, the tool will provide the compatible options included in the tool. See the available options further below.	The sequencing and library preparation reagent drop-down options will populate based on which platform is selected. If a sequencer is not feasible for the selected sample throughput within 1 year, the sequencer will not appear (e.g., iSeq for 8000 samples).

Data name	*User input requirement*	*Explanation*
Sequencing reagents	Select the sequencing reagents or kits used for the platform selected. For ONT platforms, the tool will provide the only option "(choose a bundle below)", and you will need to choose the bundle in the "Library preparation kit" cell.	The reagents selected will be auto-populated on the "Reagents and Consumables" worksheet.
Recommended number of samples per run *(for SARS-CoV-2, based on reagent kit and platform type)*	NONE	Displays the recommended number of samples per run, based on the optimal capacity of the sequencing platform and kit selected. Automatically calculated based on sequencing platform/kit capacity.[a] Used to determine how many runs are needed per week.
Number of working weeks for NGS runs per year	Enter the number of weeks the laboratory operates/will operate the sequencer.	Used to determine the number of samples per week to reach the annual throughput.
Number of samples per week to reach annual throughput	NONE	The number of samples per week needed to reach the annual throughput. Automatically calculated based on the annual throughput and working weeks for NGS. Can help determine whether the number of instruments is sufficient.
Number of sequencing run(s) needed per week	NONE The number of sequencing runs per week needed to reach the annual throughput.	Automatically calculated based on the number of working weeks, optimal number of samples per run and total number of samples. Can be helpful in determining whether platform capacity and workflow are appropriate or whether an additional instrument is warranted. If the number of sequencing runs per week is greater than 1, the tool will notify that more than one instrument may be necessary.[b]
% Sequencer loading capacity (per run)	NONE	The percentage capacity of the sequencing kit used per run. Automatically calculated based on the number of samples per week to reach the annual throughput, the number of sequencing runs per week and the recommended number of samples per run for the selected sequencing platform and sequencing reagents.[c]

Data name	*User input requirement*	*Explanation*
Library preparation kit	Select the library preparation kit used from the available options.	The kit selected will be auto-populated in the "Reagents and Consumables" worksheet.[d]
Bioinformatics	Select the bioinformatics workflow used/proposed from the drop-down menu.	Selecting from the drop-down menu will pull the necessary information from the "Bioinformatics" worksheet to calculate the bioinformatics costs in the "Data Entry and Results" worksheet.
Sequencing success rate (per run)	Enter the laboratory sequencing success rate.	The sequencing success rate should be calculated as 100%, subtracting the percentage of failures across the workflows (e.g., if the sequencing failure rate is 15%, enter 85% for the success rate).

NGS: next-generation sequencing; ONT: Oxford Nanopore Technologies; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

[a] The "Recommended number of samples per run" cell displays the number of samples necessary to fill a run for the selected sequencing kit and platform, automatically calculated to account for the size of the SARS-CoV-2 genome and the loading capacity impacts on output quality. The "Number of samples per week to reach annual throughput" is calculated based on the estimated number of samples per year and the recommended number of samples per run. If the number of samples per week to reach the annual throughput is much lower than the recommended number of samples per run, this could indicate that the laboratory does not consistently receive enough samples to fill a run in a timely manner, meaning sequencing cannot occur in real time and/or sequencing runs will be underloaded. The tool will provide "% sequencer loading capacity" as an indicator of this. The lower the percentage, the more underloaded (and less economical) the sequencing runs will be, as the number of sequencing kits required to sequence all samples will increase as the percentage of sequencer loading capacity decreases. This will drive up total sequencing costs. At this time, the tool cannot accurately capture the cost of underloaded runs. Users should note whether the percentage sequencer loading capacity is less than 75% and may need to manually increase the number of sequencing kits needed in the "Reagents and Consumables" worksheet to try to account for these increased costs.

[b] If two runs per week or higher are estimated, this will require more than one instrument to achieve high sequencing throughput and/or very efficient sequencing workflow planning.

[c] For the "% sequencer loading capacity (per run)", if the percentage ≤ 50%, a message will appear to the cell to the right stating "high reagents cost wastage – consider more samples per run ", indicating that the flow cell is not loaded to appropriate capacity. Please evaluate whether additional samples can be added or the library preparation kit changed for better, more cost-effective sequencer loading capacity.

[d] If the selected library preparation kit does not match the selected sequencing platform, a message will appear in the cell to the right stating " Change library prep kit selection to match platform in cell B16", indicating that the kit is incompatible.

Table 3 provides the drop-down menu selections for all yellow-coloured cells in the "User inputs" table. The selection in the first drop-down menu will predetermine the compatible options available in subsequent drop-down menus.

Note: Filling out the tool from top to bottom is essential to provide correct combinations and obtain accurate results.

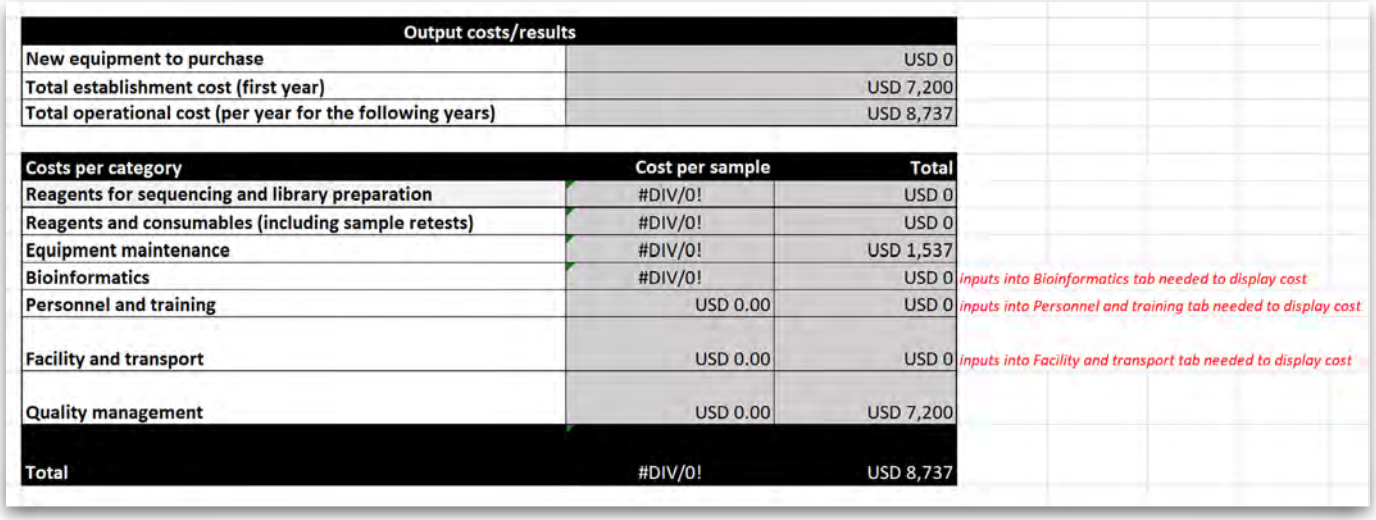
Table 3: Mapping the manufacturer and sequencing platform to the applicable library preparation kit options

MF	Sequencing platform	Sequencing Reagents	Library Preparation
Illumina	iSeq 100	• iSeq 100 i1 Reagent v2 (300-cycle) – low throughput	• EasySeq SARS-CoV-2 WGS Library Prep Kit, 96 rxns • Illumina COVIDSeq Assay, 96 samples • Illumina DNA Library Prep with tagmentation (ARTIC PCR included)
	MiniSeq	• MiniSeq High Output Reagent Kit (150 cycles) – low throughput • MiniSeq High Output Reagent Kit (300 cycles) – low throughput • MiniSeq Mid Output Reagent Kit (300 cycles) – low throughput	• EasySeq SARS-CoV-2 WGS Library Prep Kit, 96 rxns • Illumina COVIDSeq Assay, 96 samples • Illumina DNA Library Prep with tagmentation (ARTIC PCR included) • Illumina COVIDSeq Test, 3072 samples
	MiSeq	• MiSeq Reagent Kit v3 (150 cycles) • MiSeq Reagent Kit v3 (600 cycles) • MiSeq Reagent Kit v2 (300 cycles)	• EasySeq SARS-CoV-2 WGS Library Prep Kit, 96 rxns • Illumina COVIDSeq Assay, 96 samples • Illumina DNA Library Prep with tagmentation (ARTIC PCR included) • Illumina COVIDSeq Test, 3072 samples
	NextSeq 1000 NextSeq 2000	• NextSeq 1000/2000 P1 Reagents (100 cycles) • NextSeq 1000/2000 P1 Reagents (300 cycles) • NextSeq 1000/2000 P2 Reagents (100 cycles) • NextSeq 1000/2000 P2 Reagents (300 cycles) v3 – high throughput	• Illumina DNA Library Prep with tagmentation (ARTIC PCR included) • Illumina® COVIDSeq Test, 3072 samples
	NovaSeq 6000Dx	• NovaSeq 6000 SP Reagent Kit v1.5 (200 cycles) – high throughput	• Illumina DNA Library Prep with tagmentation (ARTIC PCR included) • Illumina® COVIDSeq Test, 3072 samples
ONT	MinION Mk1B MinION Mk1C GridION	Bundle options are included, select "(choose a bundle below)" option	• ONT low-cost protocol • COVID Mini: 576 samples • COVID Midi: 2304 samples • COVID Maxi: 9216 samples

MF: manufacturer; ONT: Oxford Nanopore Technologies; PCR: polymerase chain reaction; rxns: reactions.

The “Output costs/results” table (Fig. 4) is on the right of the “Data Entry and Results” worksheet. This table auto-populates based on all the user information entered in all the worksheets of the tool. Detailed information about the output costs and results will be shared towards the end of the manual.

Fig. 4. Screenshot of the “Output costs/results” table found in the “Data Entry and Results” worksheet



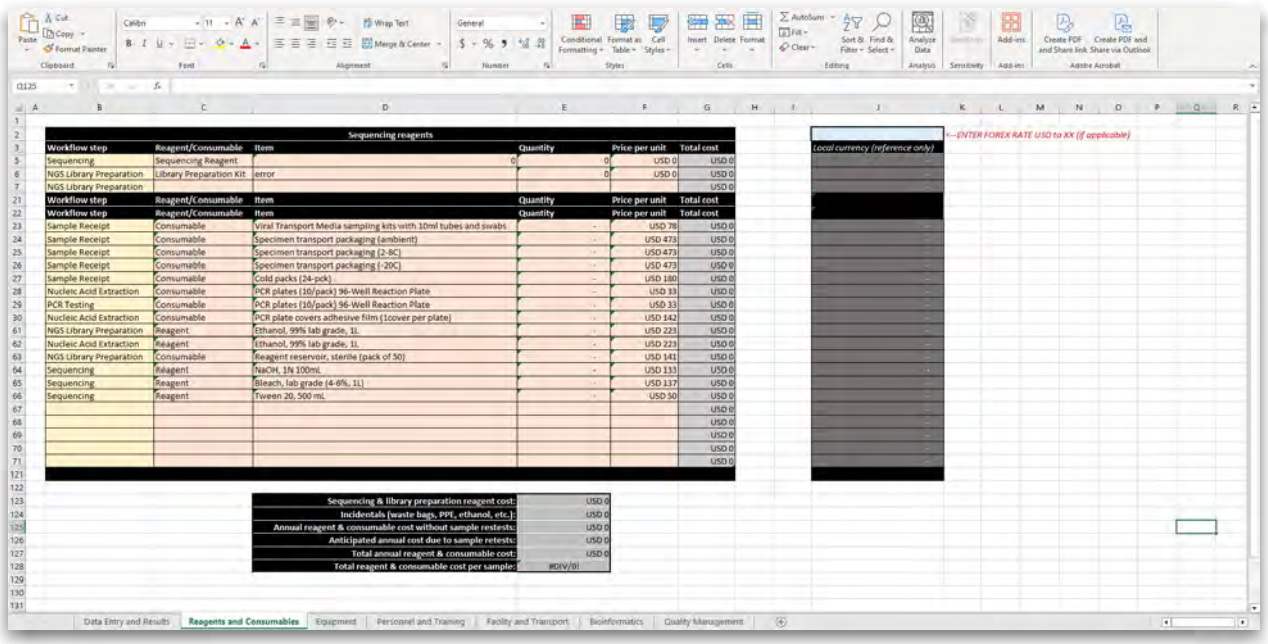
Output costs/results	
New equipment to purchase	USD 0
Total establishment cost (first year)	USD 7,200
Total operational cost (per year for the following years)	USD 8,737

Costs per category	Cost per sample	Total	
Reagents for sequencing and library preparation	#DIV/0!	USD 0	
Reagents and consumables (including sample retests)	#DIV/0!	USD 0	
Equipment maintenance	#DIV/0!	USD 1,537	
Bioinformatics	#DIV/0!	USD 0	inputs into Bioinformatics tab needed to display cost
Personnel and training	USD 0.00	USD 0	inputs into Personnel and training tab needed to display cost
Facility and transport	USD 0.00	USD 0	inputs into Facility and transport tab needed to display cost
Quality management	USD 0.00	USD 7,200	
Total	#DIV/0!	USD 8,737	

Note: Until the user enters necessary information in “Bioinformatics,” Personnel and Training,” and “Facility and Transport” worksheets, cells H15-H17 will display a message stating that inputs are required to display costs.

Reagents and consumables

Fig. 5. Screenshot of the “Reagents and Consumables” worksheet as it appears on opening the GCT[a]

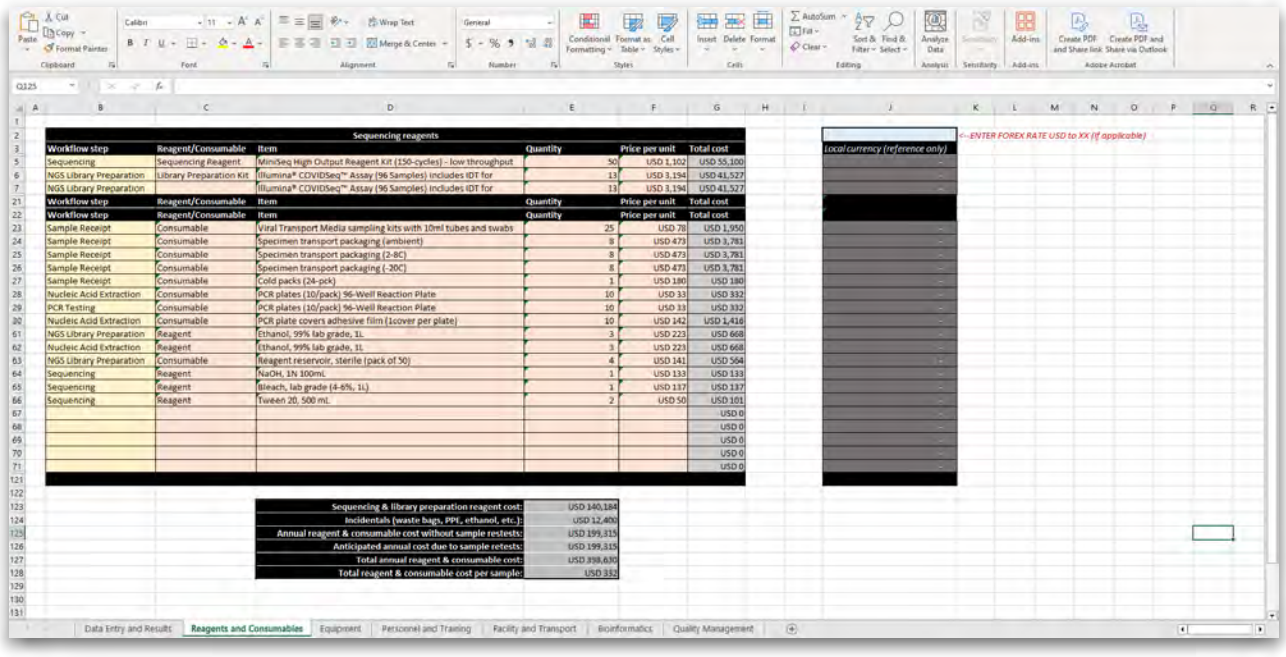


[a] Rows 8–20,31–60 and 72–120 are hidden in Fig. 5.

Once you have completed the "Data Entry and Results" worksheet, the "Reagents and Consumables" worksheet should reflect your inputs (Fig. 5). The "Reagents and Consumables" worksheet is automatically prefilled from the "User inputs" table in the "Data Entry and Results" worksheet. Each item (column D) is shown in association with the corresponding workflow step (column B) and the suggested quantity (column E) needed to process the number of samples received annually. It is important to select the appropriate workflow step for each consumable to ensure that the cost calculated per workflow step is accurate.

Fig. 6 provides an example of the auto-populated "Reagents and Consumables" worksheet following completion of "Data Entry and Results" worksheet.

Fig. 6. Screenshot of "Reagents and Consumables" worksheet following completion of "Data Entry and Results" worksheet with sample data[a]



Sequencing reagents					
Workflow step	Reagent/Consumable	Item	Quantity	Price per unit	Total cost
Sequencing	Sequencing Reagent	MiniSeq High Output Reagent Kit (150-cycles) - low throughput	50	USD 1,102	USD 55,100
NGS Library Preparation	Library Preparation Kit	Illumina® COVIDSeq™ Assay (96 Samples) includes IDT for	13	USD 3,194	USD 41,527
NGS Library Preparation		Illumina® COVIDSeq™ Assay (96 Samples) includes IDT for	13	USD 3,194	USD 41,527
Workflow step	Reagent/Consumable	Item	Quantity	Price per unit	Total cost
Workflow step	Reagent/Consumable	Item	Quantity	Price per unit	Total cost
Sample Receipt	Consumable	Viral Transport Media sampling kits with 10ml tubes and swabs	25	USD 78	USD 1,950
Sample Receipt	Consumable	Specimen transport packaging (ambient)	8	USD 473	USD 3,781
Sample Receipt	Consumable	Specimen transport packaging (2-8C)	8	USD 473	USD 3,781
Sample Receipt	Consumable	Specimen transport packaging (-20C)	8	USD 473	USD 3,781
Sample Receipt	Consumable	Cold packs (24-pck)	1	USD 180	USD 180
Nucleic Acid Extraction	Consumable	PCR plates (10/pack) 96-Well Reaction Plate	10	USD 33	USD 332
PCR Testing	Consumable	PCR plates (10/pack) 96-Well Reaction Plate	10	USD 33	USD 332
Nucleic Acid Extraction	Consumable	PCR plate covers adhesive film (1cover per plate)	10	USD 142	USD 1,416
NGS Library Preparation	Reagent	Ethanol, 99% lab grade, 1L	3	USD 223	USD 668
Nucleic Acid Extraction	Reagent	Ethanol, 99% lab grade, 1L	3	USD 223	USD 668
NGS Library Preparation	Consumable	Reagent reservoir, sterile (pack of 50)	4	USD 141	USD 564
Sequencing	Reagent	NaOH, 1N 100mL	1	USD 133	USD 133
Sequencing	Reagent	Bleach, lab grade (4-6%, 1L)	1	USD 137	USD 137
Sequencing	Reagent	Tween 20, 500 mL	2	USD 50	USD 101

Sequencing & library preparation reagent cost:	USD 140,184
Incidentals (waste bags, PPE, ethanol, etc.):	USD 12,400
Annual reagent & consumable cost without sample retests:	USD 199,315
Anticipated annual cost due to sample retests:	USD 199,315
Total annual reagent & consumable cost:	USD 398,630
Total reagent & consumable cost per sample:	USD 332

[a] Rows 8–20, 31–60 and 72–120 are hidden in Fig. 6.

The prices indicated were derived from GCT working group users and are listed in US dollars. Prefilled prices are meant to provide an orientation to the user and may not represent accurate local prices. To reflect the local procurement situation, carefully review the worksheet and adjust the pricing per unit (column F). Consider the following guidance when making changes to the "Reagents and Consumables" worksheet:

- Prices should be modified to reflect local procurement costs.
- Items can be modified to match the laboratory's workflow.
- If the laboratory procures different bulk packaging than what is suggested, the item description and quantity should be adjusted accordingly. **This will remove the formula from the cell. If a copy of this edited version of the tool is to be used for a different throughput or different costing scenario, you should begin with a new, unused file of the tool. Otherwise, this quantity will need to be updated manually.**
- Decimal numbers are not allowed.

- To add specific reagents and/or consumables to the list, use the empty rows at the bottom of the table.
- Do not add items included in the incidentals list (Table 8), as these costs are automatically calculated in the total costs.
- To exclude reagents and/or consumables from the list, indicate "0" as quantity. **Do not delete rows from the list.**
- Ensure that all items are coordinated to the correct workflow step.
- The reagents for the ONT sequencing platform consist of bulk packages, including the flow cells, library preparation kit, and barcodes required for multiplexing for multiple runs. Due to the fixed quantities, there will be a certain number of unused reactions depending on the selected throughput. However, it is typically more economical to choose these bulk items, and you may consider repurposing the remaining reactions.

A breakdown of the calculated costs for reagents and consumables can be found at the bottom of the "Sequencing reagents" table, as shown in Table 4.

Table 4: Total cost breakdown of reagents and consumables at the end of the "Sequencing reagents" input table in the "Reagents and Consumables" worksheet

Worksheet: Reagents and Consumables		
Table: Sequencing reagents		
Data name	***User input requirement***	***Explanation***
Sequencing and library preparation reagent cost	NONE Data automatically calculated	Total sum of sequencing and library preparation kit reagents and consumables listed above, based on quantities provided.
Incidentals (waste bags, PPE, ethanol, etc.)		Incidental costs include minor costs that are associated with reagents and consumables. The items included can be found in Table 5, "Incidental costs".[a]
Annual reagent and consumable cost without sample retests		Calculated annual costs to allow sequencing of the indicated number of samples when assuming a success rate of 100%.
Anticipated annual cost due to sample retests		The additional annual costs to be considered based on the "User input" sequencing success rate to reach the total number of annual throughput sample results.
Total annual reagent and consumable cost		Total annual sum of all reagents and consumables, taking into account indicated quantities, incidentals and sample retests.
Total reagent and consumable cost per sample		Breakdown of the total costs for reagents and consumables per sample, considering the indicated annual sample throughput.

[a] During the piloting of this tool, the incidental costs were validated using actual costs and therefore are reflected as a percentage (7%) of the costs related to reagents and consumables.

Table 5: Items included in the incidental cost percentage calculation

Incidental Costs
Nitrile gloves, laboratory coat/gown, tube racks, RNase AWAY, isopropanol (70%), autoclave bags (large and small), permanent markers, disinfectant/soap, paper towels, bleach or equivalent, wipes, lens paper, lab tape

Note: These incidentals should not be added to the list of reagents and consumables, as the cost is already accounted for in the tool.

Equipment

The "Equipment" worksheet will be auto-populated following completion of the "Data Entry and Results" worksheet. The "Equipment" worksheet aims to capture all laboratory equipment used for sequencing SARS-CoV-2-positive specimens. You will need to provide and review pre-populated information in this section (Fig. 7, Table 6).

Quantity: The actual quantity of each equipment item should be completed. A list of equipment to be purchased, based on the recommended quantity of each equipment item to accomplish the selected throughput on the selected platform, will be generated based on user inputs of actual equipment quantities available in the laboratory. (For example, if it is recommended that the laboratory has two PCR workstations, but only one is currently available for NGS, the "Equipment to buy" table will auto-populate, suggesting that the user needs to purchase one PCR workstation, and will indicate the estimated cost of one PCR workstation (the cost may differ for each user). **This will remove the formula from the cell. If the tool is to be used for a different throughput or different input in the "Data Entry and Results" worksheet, begin with a new, unused file of the tool. Otherwise, this quantity will need to be updated manually.**

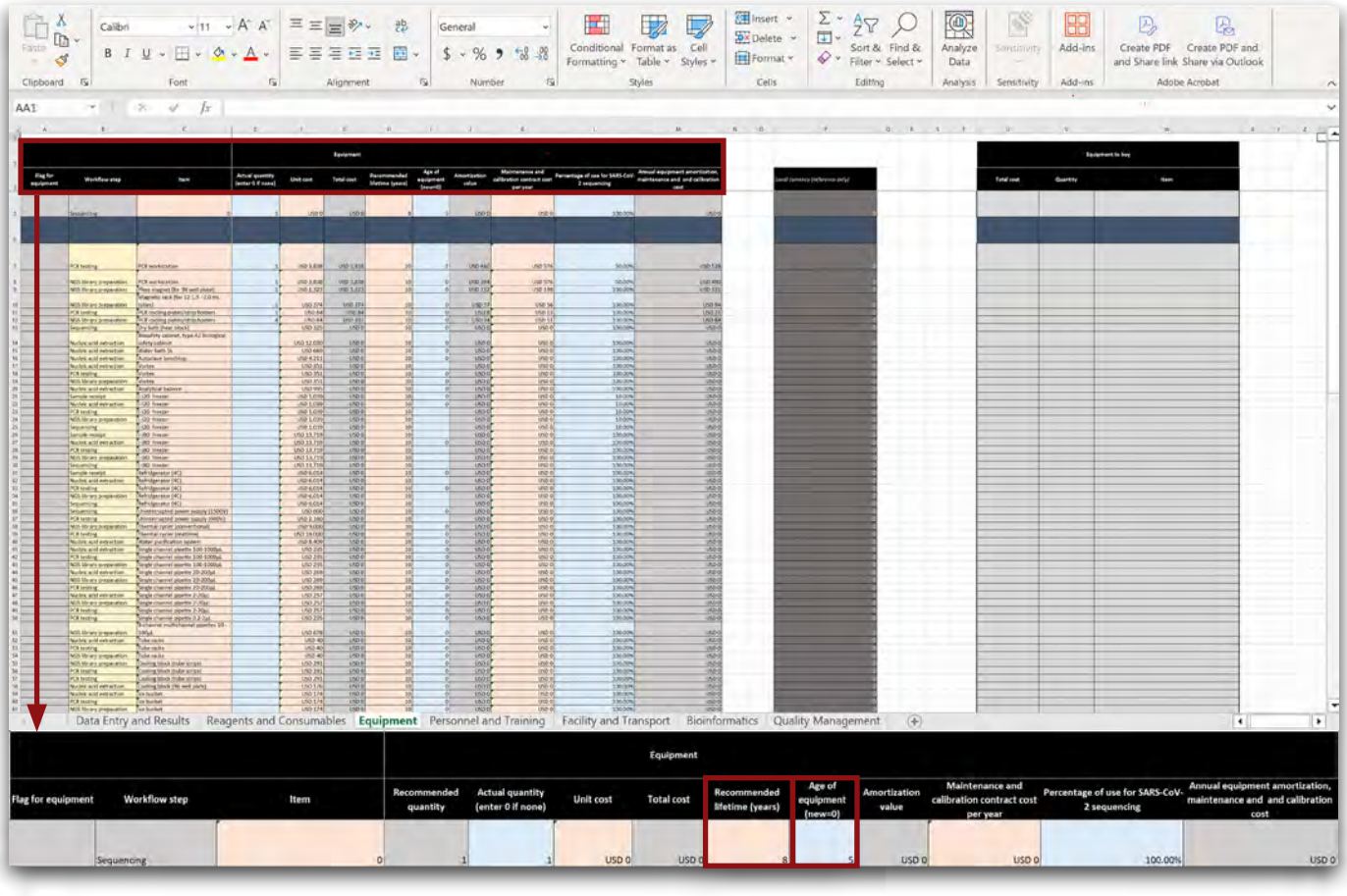
Age: The age of all equipment present in the laboratory will need to be provided; this is required for each listed item for the annual equipment costs to be considered. If the equipment is new, input the equipment age as "0". If there are several pieces of the same equipment, provide an average age. The table will also calculate the amortization value, which accounts for the depreciation of the equipment and the future need to purchase new equipment.

Table 6: Data requirements for equipment tracking in the "Equipment" worksheet

WORKSHEET: Equipment TABLE: Equipment		
Data name	***User input requirements***	***Explanation***
Note for equipment	NONE	Automatically generates the note "Buy" if the actual quantity of an item is less than the recommended quantity of the item OR if the average age of equipment is more than the recommended lifetime, suggesting that additional quantities should be purchased. The list of equipment with the note "Buy" automatically appears in the "Equipment to buy" table next to the "Equipment" table.
Workflow step	Select the workflow associated with the item.	Necessary to calculate the cost per workflow.
Item	General name of the equipment item.	If a different item is used, the name of the item can be altered.
Recommended quantity	NONE	Automatically generated based on the annual throughput of the samples selected.
Actual quantity	Input the actual quantity of the equipment present in the laboratory (enter "0" if none).	The number of sequencing instruments (of the platform type selected in the "Data Entry and Results" worksheet) used for sequencing SARS-CoV-2 samples.
Unit cost	Carefully review costs and change to reflect local procurement costs.	Pre-populated data are reference unit costs; the unit costs should be changed to reflect the actual cost of the equipment for the laboratory. Customs clearance costs should not be considered here, as they are calculated separately in the "Facility and Transport" worksheet .
Total cost	NONE	The total cost of the item is calculated using the actual quantity and the unit cost.
Recommended lifetime (years)	Carefully review the recommended lifetime of equipment and change to reflect the manufacturer lifetime recommendation.	The recommended lifetime of equipment provided is a reference. If the recommended lifetime according to the manufacturer is different than what is in the tool, input the actual equipment lifetime. This cell needs to be filled in for any new equipment that is added to the table.
Age of equipment	Input the actual age of the equipment (in years). New equipment should be given an age of "0".	If there is more than one quantity of the same equipment, input an average age of all the same equipment type. No decimals should be used here; ages should be rounded up.
Amortization value	NONE	The amortization value is automatically calculated by dividing the total cost of equipment by the number of years left for the equipment to function (recommended lifetime(years) of equipment).

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Fig. 7. Screenshot of the "Equipment" worksheet and "Equipment" table showing the cells requiring data entry for unit costs and highlighting (in the red boxes of the enlarged "Equipment" table) the recommended lifetime in years and age of equipment (new equipment should be set to "0")



Maintenance fees: The " Maintenance and calibration contract cost per year" defaults to 15% of the original equipment price. If preventative maintenance agreements or service contracts are purchased for the equipment listed, they may be amended to match actual costs. Exercise care with the sequencer and thermal cyclers, as those are typically the most expensive items in the "Equipment" table ; ensure that the "Maintenance contract cost per year" is not overinflated.

For ONT, sequencing instruments do not require preventative maintenance as the fluids are self-contained in each flow cell, which is a consumable. However, the sequencers *require* an annual L&W on each instrument, as the instruments are only leased to labs. Even if the laboratory purchases on the CapEX (capital expenditure), the L&W must be paid annually. The instrument will no longer receive updates if the laboratory does not pay this fee. If not a CapEx purchase, ONT or its distributors may remove the instrument from the laboratory if the L&W lapses by more than 60 days. The L&W fee may be entered in the "Maintenance contract cost per year" in lieu of an actual maintenance contract cost.

Sequencing use: Estimate the percent usage of each piece of equipment specifically for SARS-CoV-2 sequencing and enter this estimate into the "Percentage of use for sequencing" column. If equipment is shared with another pathogen or laboratory section, this should be considered.

For example, if a laboratory has one –20 freezer for SARS-CoV-2 that is used for sample receipt, nucleic acid extraction, PCR testing, NGS library preparation and sequencing, and it is equally shared with another pathogen sequencing program me, that would mean the total percentage used for sequencing is 50%. Since the freezer is also used for five workflow steps, the 50% would be split across all five workflows, at 10% for each. See Fig. 8.

Fig. 8. Screenshot showing an example of a single –20 freezer used equally in two pathogen NGS programmes shared evenly across the five workflow steps

Sample Receipt	(-)20 freezer	10.00%
Nucleic Acid Extraction	(-)20 freezer	10.00%
PCR Testing	(-)20 freezer	10.00%
NGS Library Preparation	(-)20 freezer	10.00%
Sequencing	(-)20 freezer	10.00%

Equipment to buy: If the laboratory has equipment procurement needs based on (1) actual quantity and recommended quantity or (2) the average age of equipment and recommended lifetime for the sample throughput and platform, the tool will automatically populate "Buy" in the "Note for equipment" column. This cost will automatically be calculated at the bottom of the worksheet in the "Equipment to buy" cell. A list of specific equipment to buy will automatically be generated in the "Equipment to buy" table next to the "Equipment" table (Fig. 9). It will include the total cost and quantity of equipment to be purchased (Fig. 10). Note that the "Equipment to buy" table will populate items in the same row as in the "Equipment=" table.

Fig. 9. Screenshot of the table for "Equipment to buy" if additional procurement needs are indicated ("Equipment" worksheet, columns U through W)

Equipment to buy		
Total cost	**Quantity**	**Item**
3837.57	1	PCR workstation
3837.57	1	PCR workstation
1323.3	1	Plate magnet (for 96 well plate)
374.133	1	Magnetic rack (for 12 1.5 - 2.0 mL tubes)
84.21	1	PCR cooling plates/strip holders
324.81	1	Dry bath (heat block)
12030	1	Biosafety cabinet, type A2 biological safety cabinet
668.868	1	Water bath 5L
4210.5	1	Autoclave benchtop
702.552	2	Vortex
702.552	2	Vortex
702.552	2	Vortex
994.881	1	Analytical balance

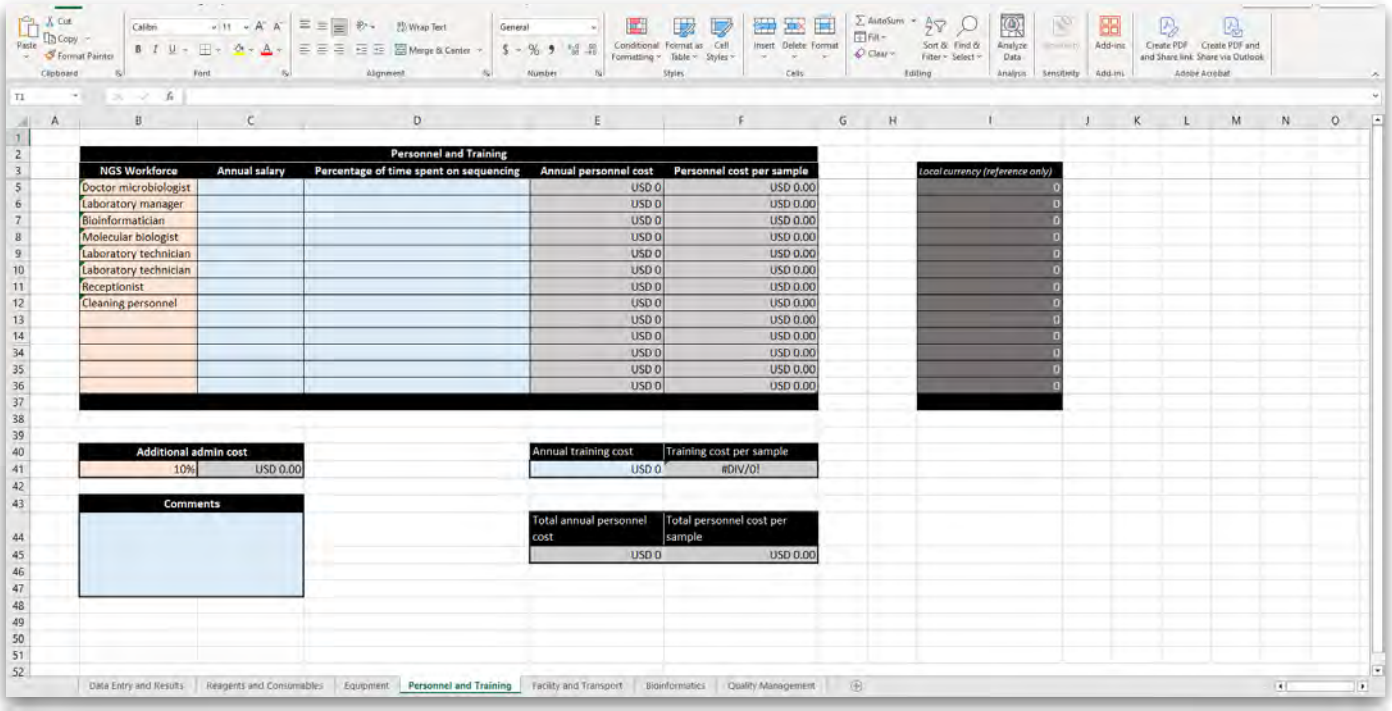
Fig. 10. Screenshot of the "Total equipment costs" table located below the "Equipment" table in columns L through H

Total recommended equipment establishment cost (including first-year maintenance and calibration)	USD 481,006.82
Equipment to buy	USD 324,180
Annual equipment amortization, maintenance and calibration cost	USD 25,704
Equipment cost per sample	USD 21.42

Personnel and training

This worksheet is used to calculate the cost of the total number of personnel involved in each workflow step of sequencing and all related personnel costs, including the cost of training (Fig. 11, Table 7).

Fig. 11. Screenshot of the "Personnel and Training" worksheet as it appears on opening the GCT[a]



[a] Rows 15-33 are hidden in Fig. 10

An additional 10% is added to the total cost of personnel to cover administrative personnel costs, such as:

- administrative and management personnel
- quality and biosafety managers
- laboratory director
- human resources personnel.

It is up to the user's discretion to alter administrative personnel costs or to add administrative personnel to the NGS workforce. When adding administrative personnel to the NGS workforce, carefully consider the percentage of time paid for by sequencing and review the prefilled "Additional admin cost percentage " cell, as it may need to be lowered if specific staff are entered into the table. In the tool, the percentage of time spent on sequencing for personnel can be adjusted. Use the actual size and profile of the laboratory to adjust the estimation. Suggested personnel are added for consideration, but are meant to function as a guide, rather than a requirement.

Table 7: Data requirements and data output for the "Personnel and Training" worksheet

WORKSHEET: Personnel and Training **TABLE: Personnel and training**		
Data input name	***User input requirements***	***Explanation***
NGS workforce	Enter all the personnel involved in sequencing SARS-CoV-2.	Example staff have been auto-populated as common workforce for consideration. Remove staff that are not used by the laboratory and add staff by using empty rows within the table.
Annual salary	Enter the annual salaries.	
Percentage of time spent on sequencing	Enter the percentage of time each member of staff spends specifically sequencing SARS-CoV-2.	E.g., if a staff member spends 30% of the time on SARS-CoV-2 NGS and 70% of the time on PCR, enter 30% in the cell.
Annual personnel cost	NONE	Automatically calculates the annual personnel cost related to sequencing based on the annual salary and the percentage of time spent on sequencing SARS-CoV-2 for each personnel member.
Personnel cost per sample	NONE	Displays the personnel cost per sample based on the annual throughput selected.
Additional admin cost (percentage)	Review administrative cost and change to reflect laboratory administration costs.	An additional 10% is recommended to cover administrative personnel costs. Consider carefully before altering this percentage.
Additional admin cost (amount)	NONE	Automatically calculated based on additional administrative cost percentage and salaries entered for personnel.

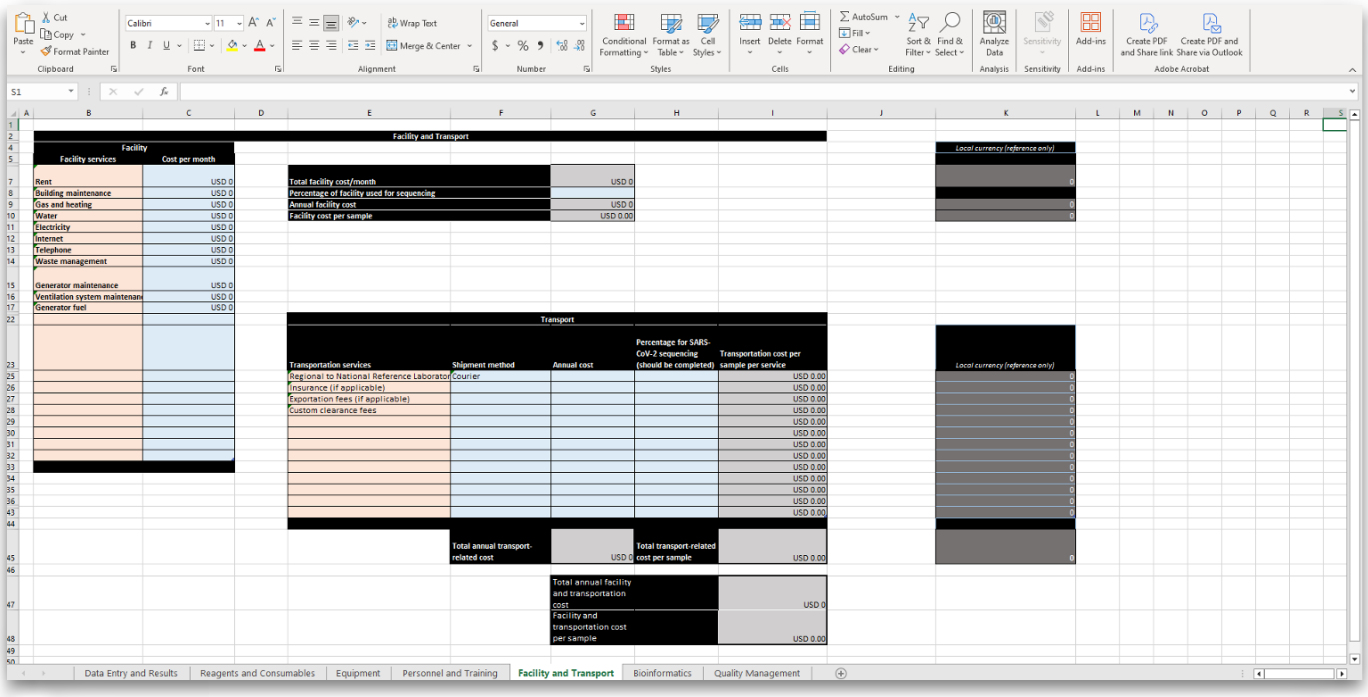
Data input name	User input requirements	Explanation
Annual training cost	Enter total annual training-related costs for all personnel specific to NGS for SARS-CoV-2.	E.g., NGS wet-lab training costs, bioinformatics training costs.
Training cost per sample	NONE	The training cost per sample based on the annual throughput.
Total annual personnel cost	NONE	The total annual personnel cost, including salaries and administrative and training costs.
Total personnel cost per sample	NONE	Total personnel and training-related costs per sample.
Comments	Use this cell to add comments.	

NGS: next-generation sequencing; PCR: polymerase chain reaction; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Facility and transport

This worksheet is used to capture and calculate all facility and transportation costs related to sequencing (Fig. 12). Costs for this section may need to be provided by the accounting or procurement office. Facility- and transportation-related costs should be included only if they apply to sequencing. Prefilled facility and transportation services are meant to serve as a guide and should be changed or added to reflect the services used by the lab.

Fig. 12. Screenshot of the "Facility and Transport" worksheet as it appears on opening the GCT[a]



[a] Rows 18–21 and 37–42 are hidden in Fig. 12

Table 8: Data requirements and explanations to calculate the necessary transportation and facility costs per sample

WORKSHEET: Facility and transport			
Table name	***Data input name***	***User input requirements***	***Explanation***
Facility	**Facility services**	Carefully review facility services and change to reflect local facility-related services.	E.g., rent, building maintenance, internet.
	Cost per month	Enter the cost per month for each service.	
	Total facility cost/ month	NONE	Indicates total facility-related costs per month.
	Percentage of facility used for sequencing	Enter the percentage of the facility used for sequencing SARS-CoV-2.	Enter the percentage of the facility used specifically for SARS-CoV-2 sequencing (e.g., if the whole laboratory processes 10 000 tests per year, and 1200 are SARS-CoV-2 sequencing tests, the number entered should be 12%).
	Annual facility cost	NONE	This is a protected cell. It shows the annual facility-related cost.
	Facility cost per sample	NONE	This is a protected cell. It shows the facility-related cost per sample
Transport	**Transportation services**	Enter all transportation-related services.	E.g., shipment of samples or sample collection kits, insurance, custom clearance for reagents.
	Shipment method	Enter the shipment method for shipment of samples.	E.g., courier company.
	Annual cost	Enter the annual cost per service.	
	Percentage for SARS-CoV-2 sequencing (should be completed)	Enter the percentage of service used specifically for SARS-CoV-2 sequencing.	
	Transportation cost per sample per service	NONE	Transportation-related cost per service per sample.
	Total annual transport-related cost	NONE	Total annual transport-related cost for all samples.
	Total transport-related cost per sample	NONE	Total transportation -related cost per sample.

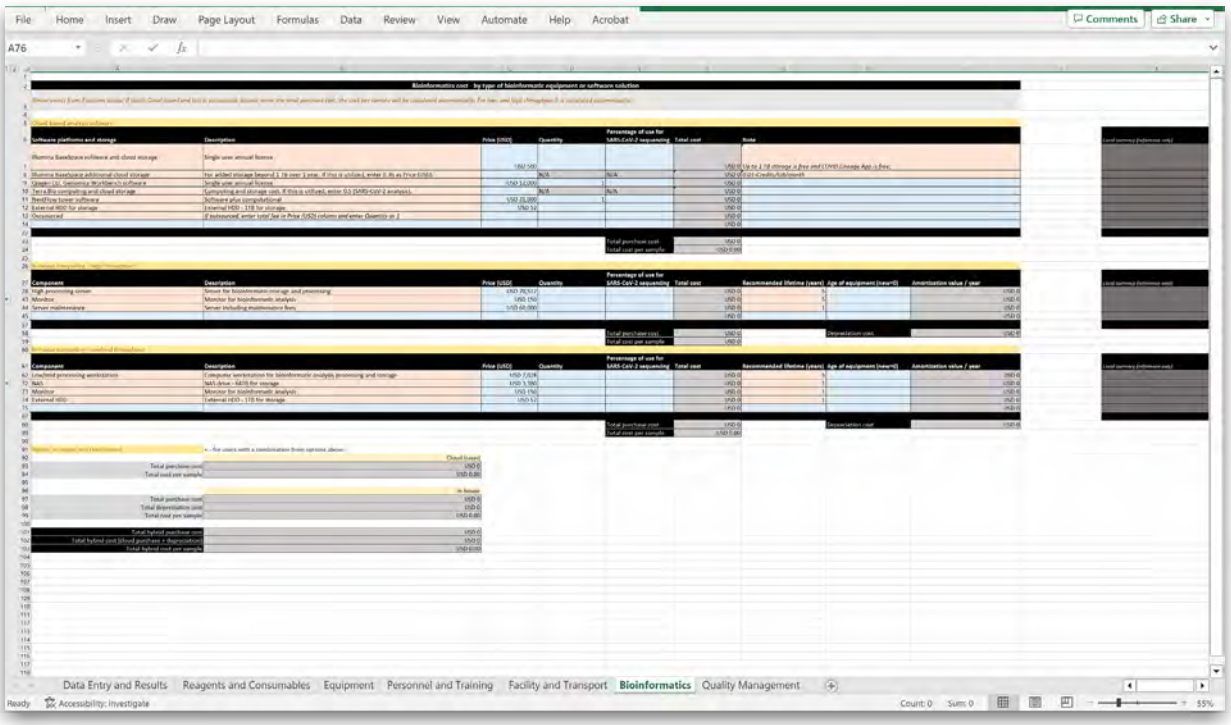
Table name	Data input name	User input requirements	Explanation
Facility and Transport totals	**Total annual facility and transportation cost**	NONE	Total annual facility- and transportation-related costs.
	Facility and transportation cost per sample	NONE	This is a protected cell. It shows the total facility- and transportation-related costs per sample.

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Bioinformatics

This worksheet calculates bioinformatics costs. It can accommodate cloud computing costs, in-house server costs or a hybrid of the two options (Fig. 13). All bioinformatics costs are captured (software licensing costs, storage, cloud costs, in-house computing hardware – low- and high-throughput options – and maintenance fees and costs for on-premises solutions. Costs are displayed as an annual total and are broken down per sample.

Fig. 13. Screenshot of the "Bioinformatics" worksheet as it appears on opening the GCT[a]



[a] Rows 15-21, 29-42, 46-56, 63-71 and 76-86 are hidden in Fig. 13

Cloud-based analysis software

Several cloud computing options are pre-populated, along with estimated costs. If any of the listed options are used, enter a quantity next to the item (with the exception of Terra.Bio and Illumina BaseSpace – see the note below Fig. 14 for details) and update the "Price (USD)" column (if necessary) to calculate total costs; likewise, enter any bioinformatic software used, along with the cost and quantity, and the percentage of use for SARS-CoV-2 sequencing in the table. If external devices (i.e., hard drives) are also used to store and/or transfer data, enter their cost and quantity in "External HDD for storage".

If bioinformatics is outsourced, please enter the price and quantity ("1" is recommended) in the "Outsourced" row to capture this information.

If cloud-based bioinformatics applications are not used, leave the "Quantity" cell blank ("0" or "NA" is acceptable) for all rows.

Fig. 14. Screenshot of components included in cloud-based analysis software for bioinformatics

Software platforms and storage	Description	Price (USD)	Quantity	Percentage of use for SARS-CoV-2 sequencing	Total cost	Note
Illumina BaseSpace software and cloud storage	Single user annual license	USD 500			USD 0	*Up to 1 TB storage is free and COVID Lineage App is free;*
Illumina BaseSpace additional cloud storage	For added storage beyond 1 TB over 1 year. If this is utilized, enter 0.36 as Price (USD).		N/A	N/A	USD 0	*0.03 iCredits/GB/month*
Qiagen CLC Genomics Workbench software	Single user annual license	USD 12,000		1	USD 0	
Terra.Bio computing and cloud storage	Computing and storage cost. If this is utilized, enter 0.5 (SARS-CoV-2 analysis).		N/A	N/A	USD 0	
Nextflow tower software	Software plus computational	USD 31,000		1	USD 0	
External HDD for storage	External HDD - 1TB for storage	USD 52			USD 0	
Outsourced	*If outsourced, enter total fee in Price (USD) column and enter Quantity as 1*				USD 0	
					USD 0	
				Total purchase cost	USD 0	
				Total cost per sample	USD 0.00	

Note: Terra.bio Computing and cloud storage and Illumina BaseSpace additional cloud storage have costs associated with individual sequence data processing and/or storage. Therefore, the "Quantity" column should not be altered, as the "Total cost" for each option is formula-based on the number of samples the laboratory processes annually (as indicated on the "Data Entry and Results" worksheet) and the "Price (USD)" column. If the laboratory uses either of these, it is recommended to use "0.36" as "Price (USD)" for Illumina BaseSpace's additional cloud storage and Terra.Bio's computing and cloud storage and to use "0.5" for SARS-CoV-2 specimens for the "Price (USD)" cell. In addition, for Illumina BaseSpace, use of apps (other than the COVID Lineage app) incurs additional costs.

In-house computing

Users who have any in-house computing capabilities should update either the "high throughput" or the "low/mid throughput" tables as described in this section . You should account for maintenance fees, redundancy (if chosen for data storage) and human resources (e.g., system administrator, systems engineer) for this on-premises option. Human resources should be captured in the "Personnel and Training" worksheet. The "Amortization value" accounts for depreciation of equipment captured in the "Bioinformatics" worksheet.

High throughput: The tool provides an option for a high-processing server; all suggested components are itemized in and can be expanded by clicking on the "+" in the left-hand pane of the Excel sheet (Fig. 15, 16). Entering a quantity in the "High processing server" "Quantity" cell will automatically update quantities of individual components and total cost (Fig. 17). The price may be adjusted to match the actual server cost. Input the percentage of use for SARS-CoV-2 sequencing. Any additional bioinformatics equipment that has or will be purchased (i.e., monitor or service maintenance fees) may be added in the "Server including maintenance fees" "Quantity" cell and appropriate quantities and prices entered.

Fig. 15. Screenshot of components included in in-house computing with high throughput

Component	Description	Price (USD)	Quantity	Percentage of use for SARS-CoV-2 sequencing	Total cost	Recommended lifetime (years)	Age of equipment (new=0)	Amortization value / year
High processing server	Server for bioinformatic storage and processing	USD 70,512			USD 0	5		USD 0
Monitor	Monitor for bioinformatic analysis	USD 150			USD 0	5		USD 0
Server maintenance	Server including maintenance fees	USD 60,000			USD 0	1		USD 0
					USD 0			USD 0
				Total purchase cost	USD 0		Depreciation cost	USD 0
				Total cost per sample	USD 0			

Fig. 16. Screenshot showing the "+" button available to display included components of the high-processing server for in-house computing with high throughput

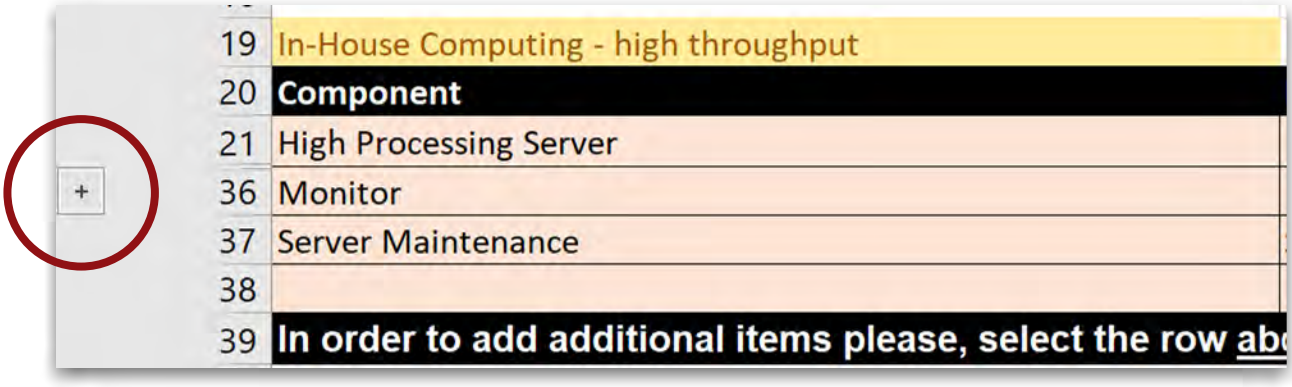
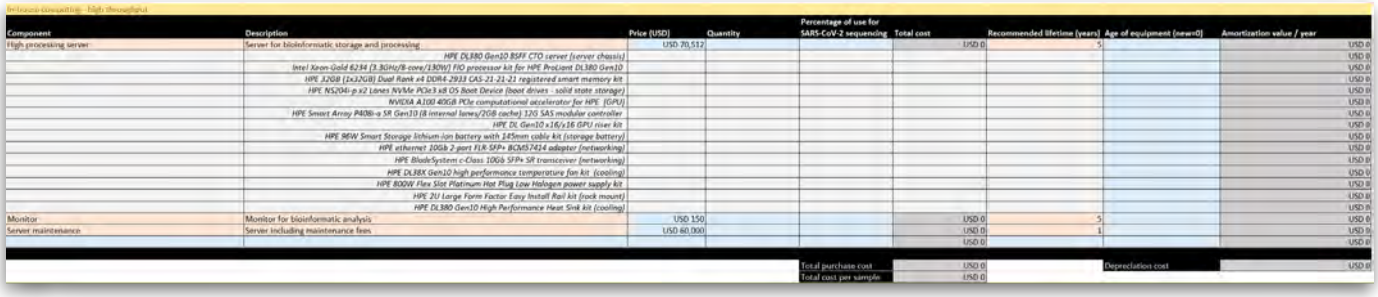


Fig. 17. Screenshot of all included components of the high-processing server for in-house computing with high throughput



Component	Description	Price (USD)	Quantity	Percentage of use for SARS-CoV-2 sequencing	Total cost	Recommended lifetime (years)	Age of equipment (new=0)	Amortization value / year
High processing server	Server for bioinformatic storage and processing	USD 70,512			USD 0	5		USD 0
	HPE DL380 Gen10 8SFF CTO server (server chassis)							USD 0
	Intel Xeon-Gold 6234 (3.3GHz/8-core/130W) FIO processor kit for HPE ProLiant DL380 Gen10							USD 0
	HPE 32GB (1x32GB) Dual Rank x4 DDR4-2933 CAS-21-21-21 registered smart memory kit							USD 0
	HPE NS204i-p x2 Lanes NVMe PCIe3 x8 OS Boot Device (boot drives - solid state storage)							USD 0
	NVIDIA A100 40GB PCIe computational accelerator for HPE (GPU)							USD 0
	HPE Smart Array P408i-a SR Gen10 (8 internal lanes/2GB cache) 12G SAS modular controller							USD 0
	HPE DL Gen10 x16/x16 GPU riser kit							USD 0
	HPE 96W Smart Storage lithium-ion battery with 145mm cable kit (storage battery)							USD 0
	HPE ethernet 10Gb 2-port FLR-SFP+ BCM57414 adapter (networking)							USD 0
	HPE BladeSystem c-Class 10Gb SFP+ SR transceiver (networking)							USD 0
	HPE DL38X Gen10 high performance temperature fan kit (cooling)							USD 0
	HPE 800W Flex Slot Platinum Hot Plug Low Halogen power supply kit							USD 0
	HPE 2U Large Form Factor Easy Install Rail kit (rack mount)							USD 0
	HPE DL380 Gen10 High Performance Heat Sink kit (cooling)							USD 0
Monitor	Monitor for bioinformatic analysis	USD 150			USD 0	5		USD 0
Server maintenance	Server including maintenance fees	USD 60,000			USD 0	1		USD 0
					USD 0			USD 0
				Total purchase cost	USD 0		Depreciation cost	USD 0
				Total cost per sample	USD 0			

Low/mid throughput: Provides an option for a low- to mid-processing server; all suggested components of the server are itemized and can be viewed by clicking on the "+" on the left-hand pane of the Excel sheet to expand the rows (Fig. 18, 19). Entering a quantity in the "Low/mid processing workstation" "Quantity" cell will automatically update quantities of individual components and total cost (Fig. 20). The price may be adjusted to match the actual server cost. Input the percentage of use for SARS-CoV-2 sequencing. Any additional bioinformatic equipment that has been or will be purchased (i.e., monitor or service maintenance fees) may be added in rows below the "Low/mid processing workstation" row , and appropriate quantities and prices entered.

Fig. 18. Screenshot of components included for in-house computing with low/mid throughput

Component	Description	Price (USD)	Quantity	Percentage of use for SARS-CoV-2 sequencing	Total cost	Recommended lifetime (years)	Age of equipment (new=0)	Amortization value / year
Low/mid processing workstation	Computer workstation for bioinformatic analysis processing and storage	USD 7,028			USD 0	5		USD 0
NAS	NAS drive - 64TB for storage	USD 3,380			USD 0	1		USD 0
Monitor	Monitor for bioinformatic analysis	USD 150			USD 0	1		USD 0
External HDD	External HDD - 1TB for storage	USD 52			USD 0	1		USD 0
					USD 0			USD 0
				Total purchase cost	USD 0		Depreciation cost	USD 0
				Total cost per sample	USD 0.00			

Fig. 19. Screenshot showing the "+" button available to display included components of the low/mid processing workstation for in-house computing with low/mid throughput

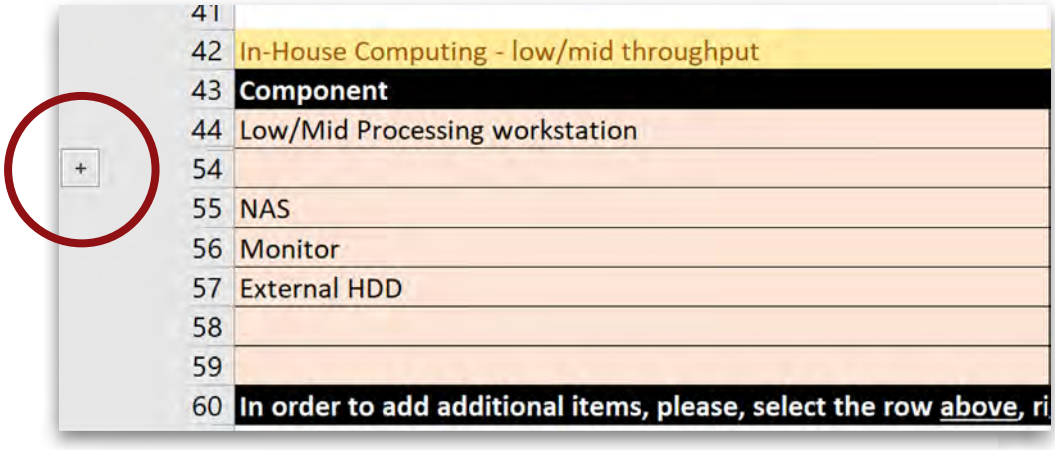
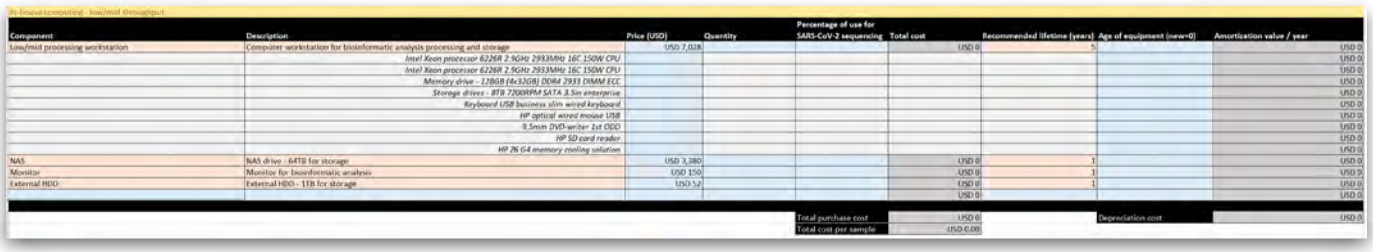


Fig. 20. Screenshot of all included components of the low/mid processing workstation for in-house computing with low/mid throughput

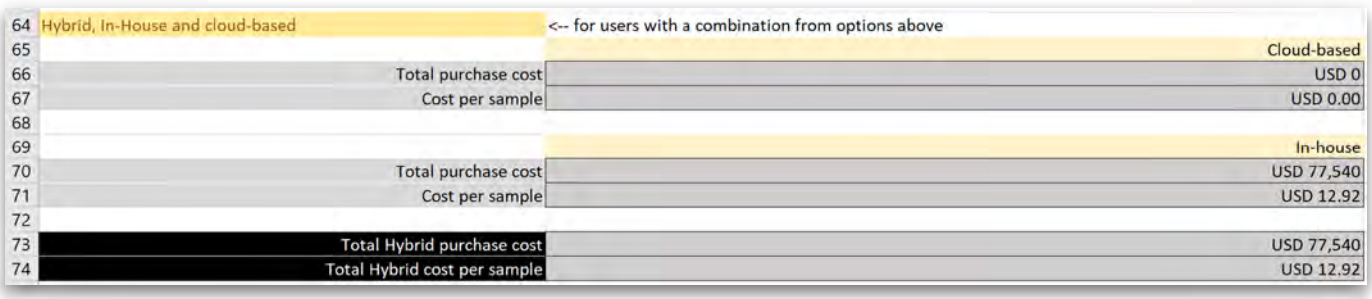


Component	Description	Price (USD)	Quantity	Percentage of use for SARS-CoV-2 sequencing	Total cost	Recommended lifetime (years)	Age of equipment (new=0)	Amortization value / year
Low/mid processing workstation	Computer workstation for bioinformatic analysis processing and storage	USD 7,028			USD 0	5		USD 0
	Intel Xeon processor 6226R 2.9GHz 2933MHz 16C 150W CPU							USD 0
	Intel Xeon processor 6226R 2.9GHz 2933MHz 16C 150W CPU							USD 0
	Memory drive - 128GB (4x32GB) DDR4 2933 DIMM ECC							USD 0
	Storage drives - 8TB 7200RPM SATA 3.5in enterprise							USD 0
	Keyboard USB business slim wired keyboard							USD 0
	HP optical wired mouse USB							USD 0
	9.5mm DVD-writer 1x ODD							USD 0
	HP SD card reader							USD 0
	HP Z6 G4 memory cooling solution							USD 0
NAS	NAS drive - 64TB for storage	USD 3,380			USD 0	1		USD 0
Monitor	Monitor for bioinformatic analysis	USD 150			USD 0	1		USD 0
External HDD	External HDD - 11B for storage	USD 52			USD 0	1		USD 0
					USD 0			USD 0
				Total purchase cost	USD 0		Depreciation cost	USD 0
				Total cost per sample	USD 0.00			

Hybrid, in-house and cloud-based

If the laboratory uses, or intends to use, both cloud-based software and in-house servers, both the "Cloud-based analysis software" and the appropriate in-house computing (either high or low/mid) sections should be completed. The workbook will automatically combine costs from both sections to give the "Total purchase cost" for both "Cloud-based" and "In-house" as well as the "Cost per sample" (Fig. 21).

Fig. 21. Screenshot of hybrid, in-house and cloud-based bioinformatics costs



64	Hybrid, In-House and cloud-based	<-- for users with a combination from options above
65		Cloud-based
66	Total purchase cost	USD 0
67	Cost per sample	USD 0.00
68		
69		In-house
70	Total purchase cost	USD 77,540
71	Cost per sample	USD 12.92
72		
73	Total Hybrid purchase cost	USD 77,540
74	Total Hybrid cost per sample	USD 12.92

Quality management

This worksheet is used to calculate the quality management costs related to sequencing SARS-CoV-2, which can include but are not limited to accreditation, proficiency testing panels, certification and other costs (Fig. 22, Table 9).

Fig. 22. Screenshot of the "Quality Management" worksheet as it appears on opening the GCT

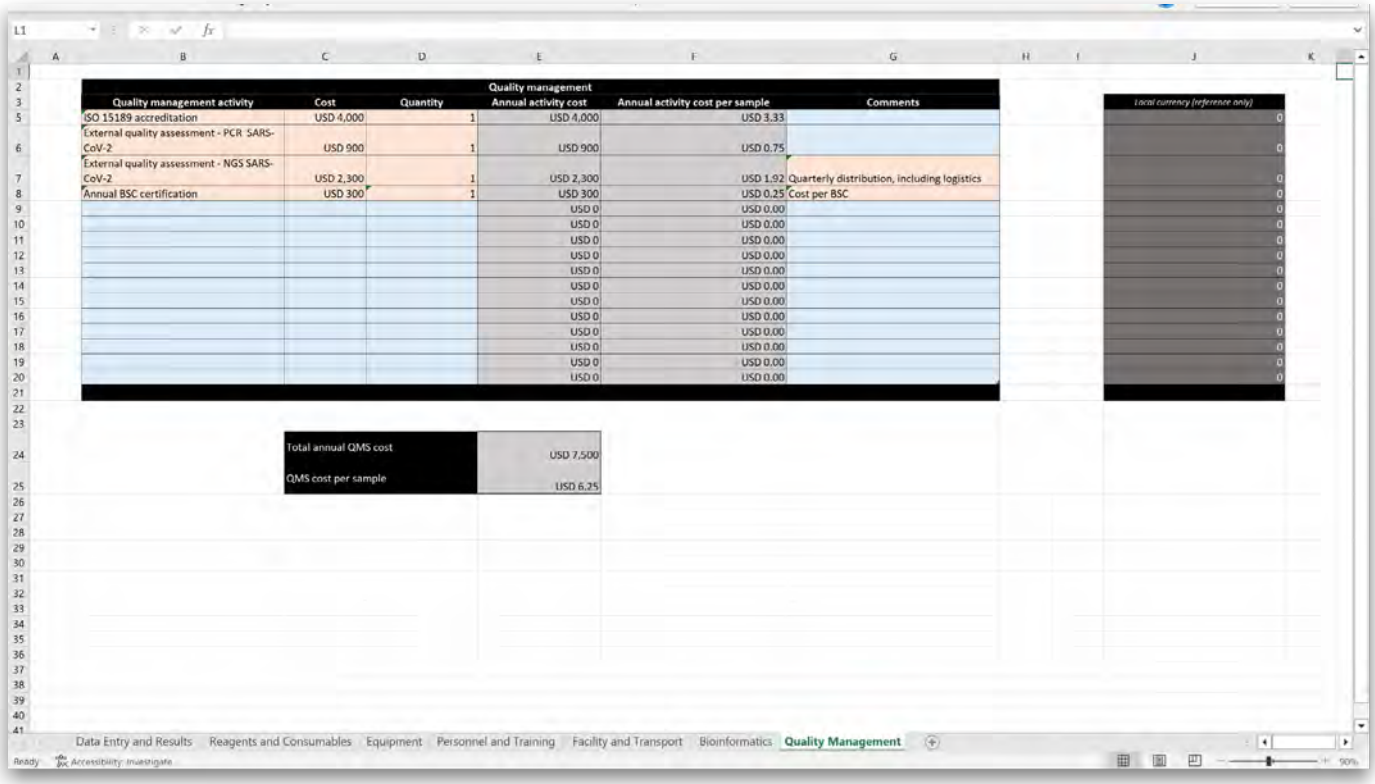


Table 9: Data requirements and explanations to calculate the cost of a QMS for genomics capacity

Worksheet: Quality Management **Table: Quality management**		
Data input name	***User input requirements***	***Explanation***
Quality management activity	Enter all quality management activities related to sequencing SARS-CoV-2.	E.g., proficiency testing panel for SARS-CoV-2 NGS.
Cost	Enter the cost for each activity.	
Quantity	Enter the quantity of each activity within a year.	E.g., if pipette calibration occurs every 6 months, enter "2".
Annual activity cost	NONE	Total annual cost for the activity automatically calculated based on activity cost and quantity.
Annual activity cost per sample	NONE	The cost of the activity per sample based on the annual throughput.

Data input name	*User input requirements*	*Explanation*
Comments	Add any comments	E.g., month of annual BSC certification or date of ISO accreditation granted.
Total annual QMS cost	NONE	Total annual cost for all QMS activities.
QMS cost per sample	NONE	Cost for all QMS activities per sample.

BSC: biosafety cabinet; NGS: next-generation sequencing; QMS: quality management system; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Costing output/results

After you have completed the "User inputs" table in the "Data Entry and Results" worksheet, and all subsequent worksheets, the "Output costs/results" table will be populated with costing information in the "Data Entry and Results" worksheet. The data in this part of the worksheet are protected and cannot be changed by the user. This section will review all the costing information output by the tool (Table 10, Table 11).

Table 10: Detailed explanation of each cell in the "Output costs/ results" table in the "Data Entry and Results" worksheet

WORKSHEET: Data Entry and Results TABLE: Output Costs/Results		
Data name	*User input requirement*	*Explanation*
New equipment to purchase	NONE Data automatically calculated after all worksheets are complete	The total cost of equipment the laboratory needs to purchase to complete the annual throughput. This cost considers the quantity and age of the equipment at the laboratory compared to the recommended lifespan of the equipment and the recommended quantity of each piece of equipment.
Total establishment cost (first year)		The total cost to establish and run a sequencing laboratory for 1 year, including purchasing new equipment.
Total operation cost (per year for the following years)		The total cost to run the laboratory for each year following the year of establishment.

Table 11: Detailed explanation of the "Costs per category" table in the "Data Entry and Results" worksheet

WORKSHEET: Data Entry and Results TABLE: Costs per category		
Data name	***User input requirement***	***Explanation***
Reagents for sequencing and library preparation	NONE Data automatically calculated after all worksheets are complete	The total cost of sequencing and library preparation reagents and consumables per sample and per year
Reagents and consumables *(including sample retests)*		The total cost of all reagents and consumables per sample and per year, including sample retests due to sequencing failure
Equipment maintenance		The total cost of equipment maintenance per sample and per year
Bioinformatics		The total cost of bioinformatics per sample and per year
Personnel and training		The total cost of personnel and training per sample and per year
Facility and transport		The total cost of facility and transport per sample and per year
Quality management		The total cost of quality management per sample and per year
Total		The total cost of running the laboratory per sample and per year for the selected options in the "User inputs" table and the other six worksheets

Table 12 shows the costs per lab workflow step per sample and the total cost per workflow step as indicated by the user, with drop-down selections for workflow in the "Reagents and Consumables" and "Equipment" worksheets. This might be helpful if different workflows are performed by different stakeholders, or to see what portions of the workflow are more costly or where efficiencies in the workflow might be helpful.

Table 12: Detailed explanation of the "Costs per workflow step" table, both per sample and the total cost

WORKSHEET: Data Entry and Results TABLE: Costs per lab workflow step		
Data name	***User input requirement***	***Explanation***
Sample receipt	NONE Data automatically calculated after all worksheets are complete	The total cost for the sample receipt workflow, provided per sample and per year
Nucleic acid extraction		The total cost for the nucleic acid extraction workflow, provided per sample and per year
PCR testing		The total cost for the PCR testing workflow, provided per sample and per year
NGS library preparation		The total cost for the NGS library preparation workflow, provided per sample and per year
Sequencing		The total cost for the sequencing workflow, provided per sample and per year
Bioinformatics		The total cost for the bioinformatics workflow, provided per sample and per year
Total		The total cost of running the laboratory per year for the selected options in the "User inputs" table and the other six worksheets.

Note: The costs for personnel and training, facility and transport, quality management, sample retest and incidentals are divided evenly among the six workflows for easier calculation. Refer to the "Reagents and consumables" section for more information on retests and incidentals.

Note: All workflow costs considered in this tool are only for sequencing SARS-CoV-2-positive specimens. For example, if the laboratory receives 10 000 SARS-CoV-2 samples per year for PCR diagnostic testing and 600 positive samples are sent for sequencing, the PCR testing workflow will account for PCR testing of 600 samples.

Limitations of the tool

While the GCT working group strove to make this tool as robust as possible to evaluate the true cost of SARS-CoV-2 sequencing programmes , the tool does have limitations that should be noted.

- This tool is specific for SARS-CoV-2. Future versions of the tool will be able to cost for other pathogens, pathogen-agnostic and metagenomic approaches, plus complex matrices like wastewater sequencing efforts.
- The current version of the tool is only able to assess and cost a single sequencing platform and instrument model. It cannot evaluate labs that may have mixed sequencing platforms (i.e., a lab with an Illumina MiSeq and a NextSeq 1000; or a lab with an ONT GridION and an Illumina MiniSeq). In future versions of the tool, users will be able to capture multiple systems from various manufacturers. To generate the costs for multiple systems, you can complete this tool multiple times and zero out any items that were captured in the first iteration that do not need to be counted multiple times (e.g., “Personnel and Training”, “Facility and Transport”, “Bioinformatics” and “Quality Management”), since this information will have been captured on the first completed version of the tool for the laboratory.
- Reagents available for selection in the tool are limited and are paired with each sequencing platform. While other reagents (library preparation kits) can be used, the items selected for the GCT are the kits most paired with the sequencing platforms. You can manually change the prefilled sequencing library preparation kit names, quantities and prices to reflect the reagents used within the laboratory. It is important to note that formulas in peach-coloured cells will be deleted. To cost a different scenario, begin with a new, unused copy of the tool.

Annex 1. Scenarios for GCT use

This annex lists four scenarios for which this tool may be used. The scenarios do not represent a comprehensive list of situations where the tool may be useful but rather are intended as a guide. Listed in each scenario is the objective, the desired purpose of using the tool, the required inputs, what information is needed from the user to fulfil the purpose of using the tool and the expected outcome.

Scenario #1: Validate	
Objective	I wish to validate a recent costing exercise conducted at my lab for NGS activities.
User required inputs	• Provide the same inputs to the GCT as provided for the recent costing exercise completed by the laboratory. • Carefully complete all portions of the tool to account for the total cost of supporting SARS-CoV-2 NGS – including inputting actual package sizes, quantities and cost of the item purchased by laboratory, rather than using reference package size and pricing available in the tool. • For user entries in a currency different than USD, please use the same currency conversion of the local currency to USD for all items in the GCT.
Expected outcome	The tool allows you to compare and validate a recent costing exercise completed by the laboratory to ensure accurate costing.

Scenario #2: Routine	
Objective	To determine the cost of the NGS operations occurring within my laboratory, then to routinely use the tool to stay updated on the operational costs of my NGS laboratory.
User required inputs	• Input the actual cost and quantities for all information in each worksheet. It is important that all equipment costs represent local procurement costs. • If equipment items are pre-populated within the tool, and do not exist in the laboratory, the item quantity should be changed to "0". • Review the tool annually to ensure accurate representation of reagents and consumables procurement for the annual throughput. • Update the tool to represent any changes within the laboratory as they occur (hiring new staff, pay raises, new equipment, etc.). ▪ See the "Additional considerations on the use of the tool" section for guidance on updating the tool.
Expected outcome	The tool will provide the total annual running costs and cost per sample for a laboratory operating NGS for SARS-CoV-2.

Scenario #3: Optimize	
Objective	To make informed decisions on sequencing activities by optimizing the sequencing workflows and budget for NGS.
User required inputs	• Expect to use the tool two or more times for this scenario. • Provide all accurate information for the current NGS costing situation of the laboratory. • Create a copy of the tool used to cost the current situation of the laboratory. (By using a copy, you will not need to re-enter stable information such as personnel, transportation or facility costs.) You should change certain selections in the tool that represent the changes the laboratory may make. • Repeat using the copy of the tool with different possible scenarios for the laboratory to see how the cost per sample may be optimized. • **Note: If you have altered peach-coloured cells (cells containing formulas) in the initial use of the tool, review the information in these cells during subsequent use to accurately project costing changes.**
Expected outcome	For every scenario, the tool will produce a breakdown of annual running costs as well as costs per sample that should be saved. The scenarios can be compared, and they might highlight adjustments that could be made to optimize the budget for NGS.

Scenario #4: Scale-up	
Objective	To determine the overhead costs to scale up my lab from low throughput to high throughput – size of workforce, allocation of time, package size of reagents, etc.
User required inputs	• Provide inputs for all worksheets for the current throughput of the laboratory, paying careful attention to the equipment available. • Following completion of the tool, open a new, unused copy of the tool and rework through the "User inputs" table to enter the new desired throughput. • For the "Equipment" worksheet, it is important to input the same "Actual quantity" and "Age" of the equipment specified in the initial use of the tool. Here, the tool will calculate the cost to add the necessary equipment to scale up to a higher throughput. • If possible, you may receive quotes for needed equipment and input the cost of procuring this equipment under "Unit cost" (column F) in the "Equipment" table. • Consider the workforce needed to process an increase in throughput and be sure to include the extra personnel and projected salaries in the subsequent use of the tool.
Expected outcome	Filling the tool with the current throughput of the laboratory, then comparing the results to a second use of the tool with increased throughput, will indicate what new equipment needs to be added to the laboratory to successfully support the desired higher throughput. For the "User inputs" table in the "Data Entry and Results" worksheet, the tool will indicate whether the increased throughput will fill the sequencing platform to a higher capacity, therefore using the loading capacity more efficiently and potentially leading to cost savings. In the "Reagents and Consumables" worksheet, the user may be recommended to purchase bulk packaging with the increased throughput, which may lead to cost savings. Comparing the current use of equipment and reagents to the suggestions provided for increased throughput will help the laboratory to decide whether scaling up the throughput will provide cost savings per sample, and whether scaling up is financially feasible.

GCT: Genomics Costing Tool; NGS: next-generation sequencing; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

NOTE: For any scenario, changing the peach-coloured cells will remove the formula from the cell. Consequently, you will need to go through all the worksheets and tables to ensure that all cells contain correct information. If the tool is to be used for a different throughput or different input in the "Data Entry and Results" worksheet, begin with a new, unused file of the tool or prefilled data will need to be updated manually.

References

1. Global genomic surveillance strategy for pathogens with pandemic and epidemic potential, 2022–2032. Geneva: World Health Organization; 2022 (https://iris.who.int/handle/10665/352580, accessed 27 November 2023).
2. Laboratory test costing tool. Copenhagen: World Health Organization Regional Office for Europe; 2019 (**https://iris.who.int/handle/10665/346135**, accessed 27 November 2023).

Contact

Department of Epidemic and Pandemic Preparedness and Prevention

World Health Organization
20, avenue Appia
1211 Geneva 27 Switzerland

Email: **pathogengenomics@who.int**

Website: **www.who.int/initiatives/genomic-surveillance-strategy**