



# SAMPLING STRATEGIES FOR GENOMIC SURVEILLANCE

*SOUTH & SOUTHEAST ASIA*  
*PATHOGEN GENOMICS PRIORITIZATION & IMPLEMENTATION WORKSHOP*  
*September 9-13, 2024*  
*Bangkok, Thailand*

## WORKSHOP PARTNERS



Asia Pathogen  
Genomics Initiative



CENTRE FOR  
PATHOGEN  
GENOMICS

Sydney Infectious Diseases Institute  
Centre for Infectious Diseases & Microbiology  
WHO Southeast Asia Regional Office (SEARO)  
WHO Western Pacific Regional Office (WPRO)  
WHO International Pathogen Surveillance Network (IPSN)

# Genomic Surveillance Sampling Strategies

## Outline

1. Background
2. Sampling Types
3. Sampling Strategy Early Tool Development
4. Discussion: Key Questions & Tool Needs



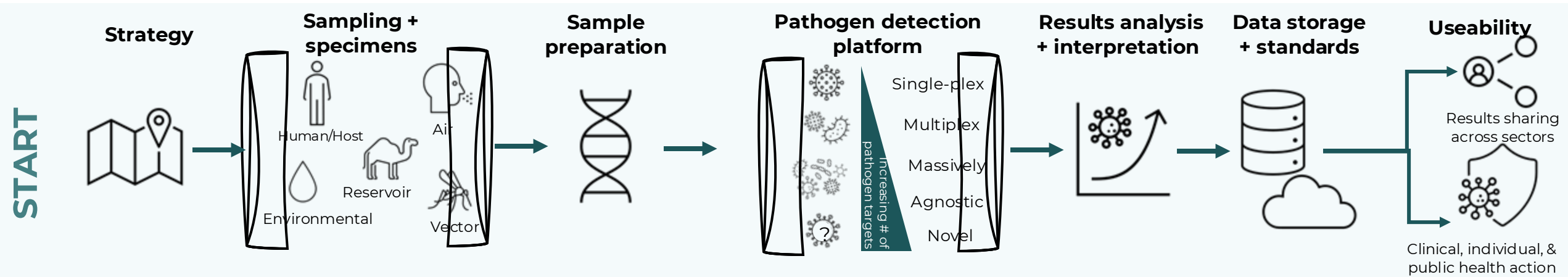


# Genomic Surveillance Sampling Strategies

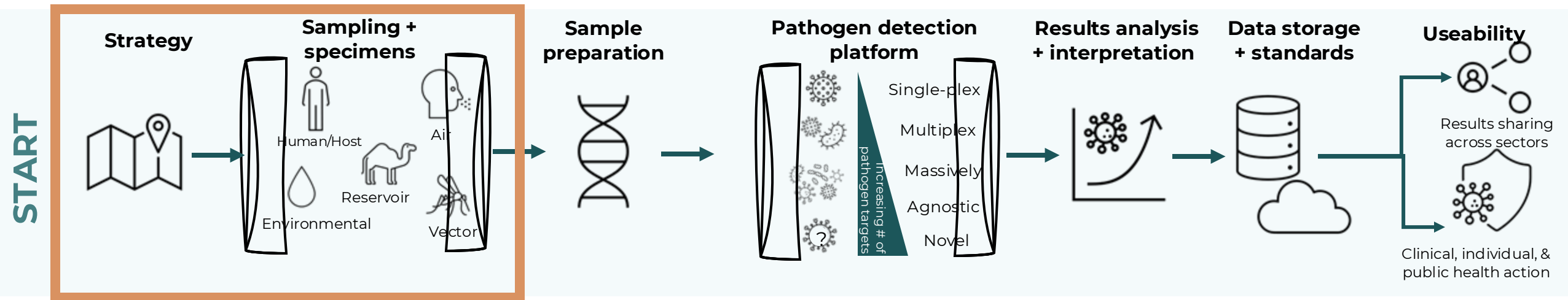
Background



# Pathogen genomics workflow



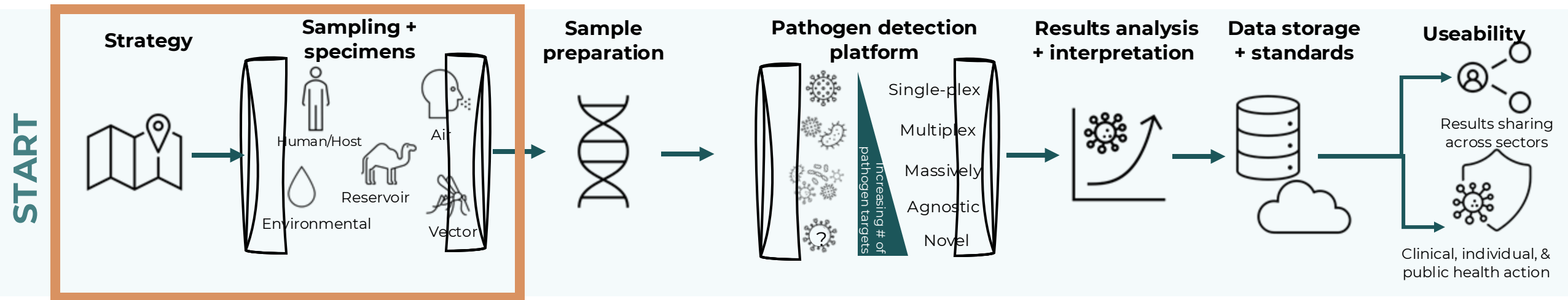
# Pathogen genomics workflow



*Sampling strategies are the foundation of the workflow...*



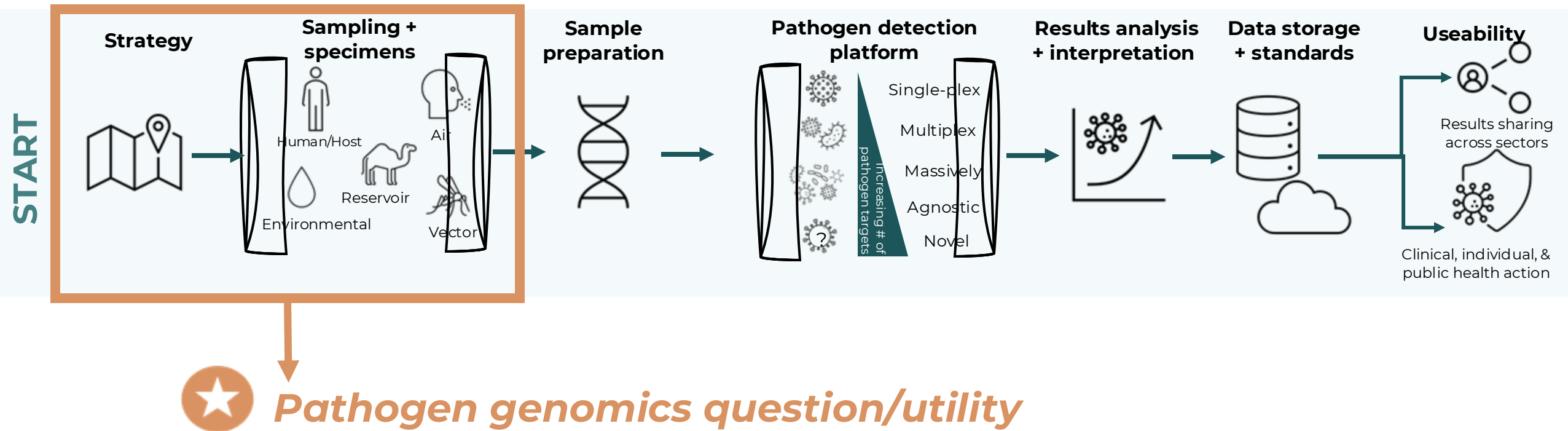
# Pathogen genomics workflow



- *How many specimens need to be sequenced to answer my question?*
- *What specimens should I prioritize during an outbreak?*
- *What are the building blocks variables critical to sampling strategies?*



# Pathogen genomics workflow



# Genomic Surveillance Sampling Strategies

## **Pathogen Genomics Purpose/Question**



Early detection and warning systems

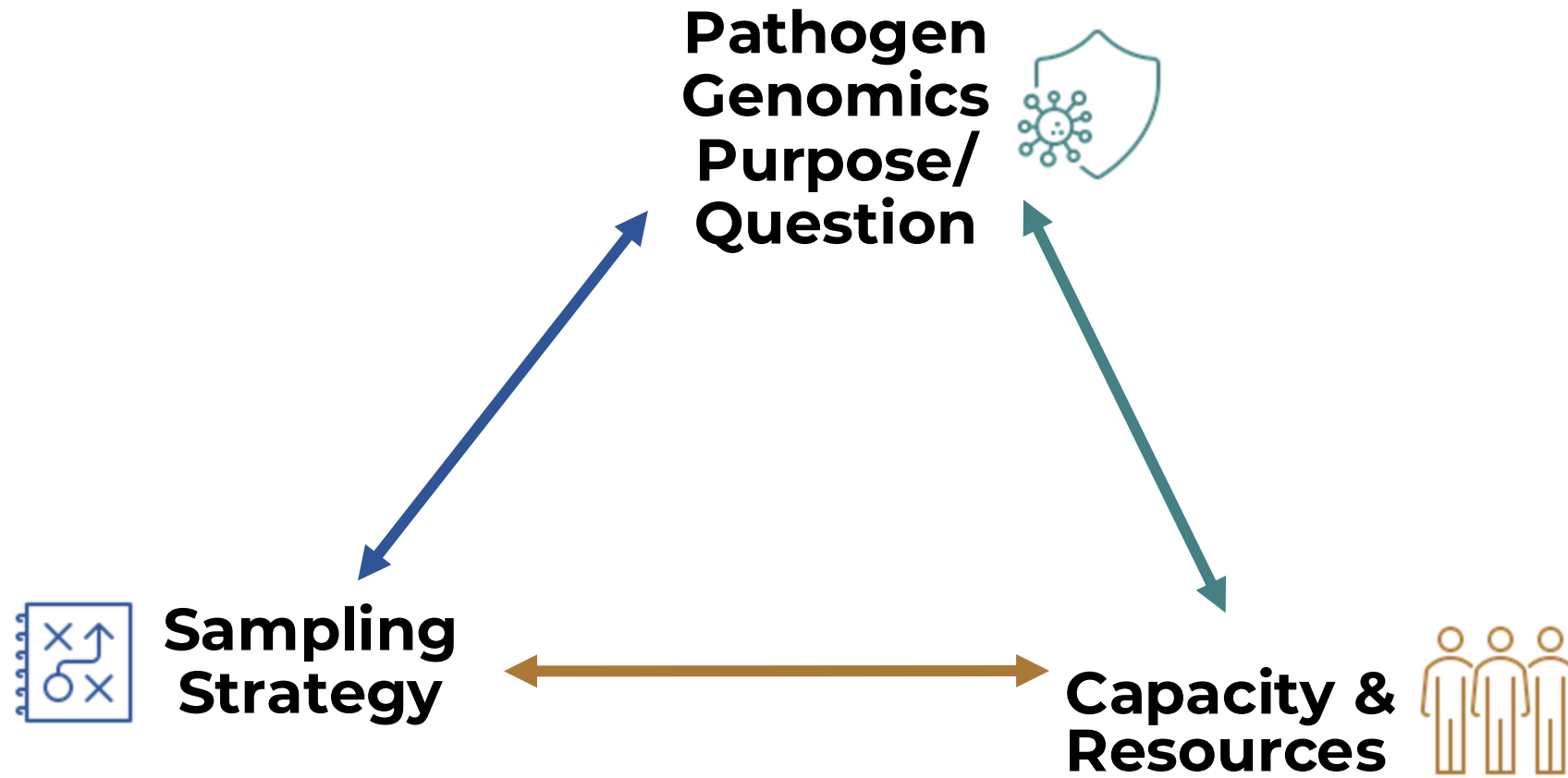
Identify and track variants, intervention development

Intervention effectiveness, transmission, and resistance





# Genomic Surveillance Sampling Strategies



# Sampling strategy building blocks

Specimen  
type

Target  
population

Sensitivity

Cost

Surveillance  
systems

Epidemiology

Specimen  
availability

Workforce

Genomics  
utility

Timeline

Capacity

Mutation rate

Other  
detection  
tools



# Sampling strategy building blocks

## Epidemiology

As **pathogen variant prevalence increases**, the general number of specimens sequenced needed likely **decreases**.

## Cost

As **sequencing workflow cost increases**, the general number of specimens able to be sequenced likely **decreases**.

## Mutation rate

As **pathogen mutation rate increases**, the general number of sequences needed likely **increases**.

## Sensitivity

As **genomic surveillance system sensitivity increases**, the general number of sequences needed likely **increases**.

## Capacity

As **sequencing capacity increases**, the general number of sequences able to be sequenced likely **increases**.

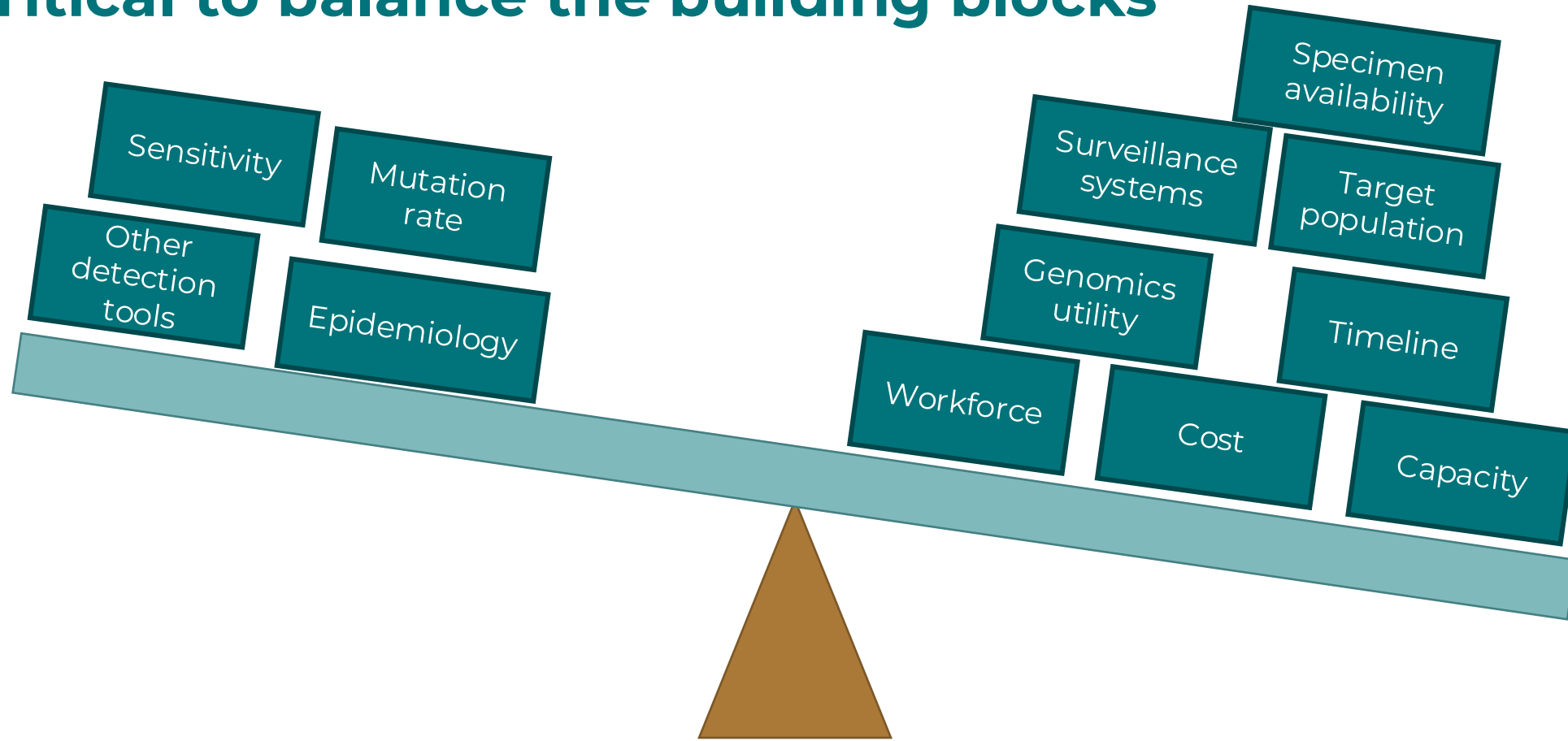
## Specimen type

As genomics surveillance system **specimens shift from individual clinical to environmental**, the general number of specimens needed likely **decreases**.



# Sampling strategy building blocks

**Critical to balance the building blocks**



# Types of sampling strategies

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



### Genomic application

- Random selection of specimens from all SAR-CoV-2 infections to estimate variant prevalence



# Types of sampling strategies

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



### Genomic application

- Random selection of specimens from all SAR-CoV-2 infections to estimate variant prevalence

## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



### Genomic application

- Dividing pediatric and adult specimens for sequencing to better understand if there are malaria genomic differences



# Types of sampling strategies

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



### Genomic application

- Random selection of specimens from all SAR-CoV-2 infections to estimate variant prevalence

## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



### Genomic application

- Dividing pediatric and adult specimens for sequencing to better understand if there are malaria genomic differences

## Cluster Sampling

### Description

- Target population is divided into clusters (e.g., geographic areas, institutions), and entire clusters are randomly selected for sequencing sampling



### Genomic application

- Randomly sampling MERS cases from different health centers with ongoing outbreaks



# Types of sampling strategies

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



### Genomic application

- Random selection of specimens from all SAR-CoV-2 infections to estimate variant prevalence

## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



### Genomic application

- Dividing pediatric and adult specimens for sequencing to better understand if there are malaria genomic differences

## Cluster Sampling

### Description

- Target population is divided into clusters (e.g., geographic areas, institutions), and entire clusters are randomly selected for sequencing sampling



### Genomic application

- Randomly sampling MERS cases from different health centers with ongoing outbreaks

## Purposive (Targeted) Sampling

### Description

- Sampling focuses on specific groups or characteristics relevant to the research question (e.g., targeting samples from regions or individuals with known exposures)



### Genomic application

- Sequencing specific severe Mpox cases to better understand if specific mutations may be related to virus severity





# Types of sampling strategies

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



### Genomic application

- Random selection of specimens from all SAR-CoV-2 infections to estimate variant prevalence

## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



### Genomic application

- Dividing pediatric and adult specimens for sequencing to better understand if there are malaria genomic differences

## Cluster Sampling

### Description

- Target population is divided into clusters (e.g., geographic areas, institutions), and entire clusters are randomly selected for sequencing sampling



### Genomic application

- Randomly sampling MERS cases from different health centers with ongoing outbreaks

## Purposive (Targeted) Sampling

### Description

- Sampling focuses on specific groups or characteristics relevant to the research question (e.g., targeting samples from regions or individuals with known exposures)



### Genomic application

- Sequencing specific severe Mpox cases to better understand if specific mutations may be related to virus severity

## Convenience Sampling

### Description

- Samples are collected based on their ease of accessibility rather than random selection; often used when resources or time are limited



### Genomic application

- Sequencing the 10 first specimens from a cholera outbreak because these were the only samples that remained on cold-chain



# Logistics may become easier and cost may decrease, but bias may increase

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



### Genomic application

- Random selection of specimens from all SAR-CoV-2 infections to estimate variant prevalence

## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



### Genomic application

- Dividing pediatric and adult specimens for sequencing to better understand if there are malaria genomic differences

## Cluster Sampling

### Description

- Target population is divided into clusters (e.g., geographic areas, institutions), and entire clusters are randomly selected for sequencing sampling



### Genomic application

- Randomly sampling MERS cases from different health centers with ongoing outbreaks

## Purposive (Targeted) Sampling

### Description

- Sampling focuses on specific groups or characteristics relevant to the research question (e.g., targeting samples from regions or individuals with known exposures)



### Genomic application

- Sequencing specific severe Mpox cases to better understand if specific mutations may be related to virus severity

## Convenience Sampling

### Description

- Samples are collected based on their ease of accessibility rather than random selection; often used when resources or time are limited



### Genomic application

- Sequencing the 10 first specimens from a cholera outbreak because these were the only samples that remained on cold-chain



# Table discussion

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



## Cluster Sampling

### Description

- Target population is divided into clusters (e.g., geographic areas, institutions), and entire clusters are randomly selected for sequencing sampling



## Purposive (Targeted) Sampling

### Description

- Sampling focuses on specific groups or characteristics relevant to the research question (e.g., targeting samples from regions or individuals with known exposures)



## Convenience Sampling

### Description

- Samples are collected based on their ease of accessibility rather than random selection; often used when resources or time are limited



**Which of these strategies have you implemented?  
What was the pathogen/use case?**



# Table discussion

## Random Sampling

### Description

- Selecting samples from a target population in a way that each individual or unit has an equal chance of being chosen



## Stratified Sampling

### Description

- Target population is divided into subgroups (strata) based on certain characteristics, and samples are randomly drawn from each stratum



## Cluster Sampling

### Description

- Target population is divided into clusters (e.g., geographic areas, institutions), and entire clusters are randomly selected for sequencing sampling



## Purposive (Targeted) Sampling

### Description

- Sampling focuses on specific groups or characteristics relevant to the research question (e.g., targeting samples from regions or individuals with known exposures)



## Convenience Sampling

### Description

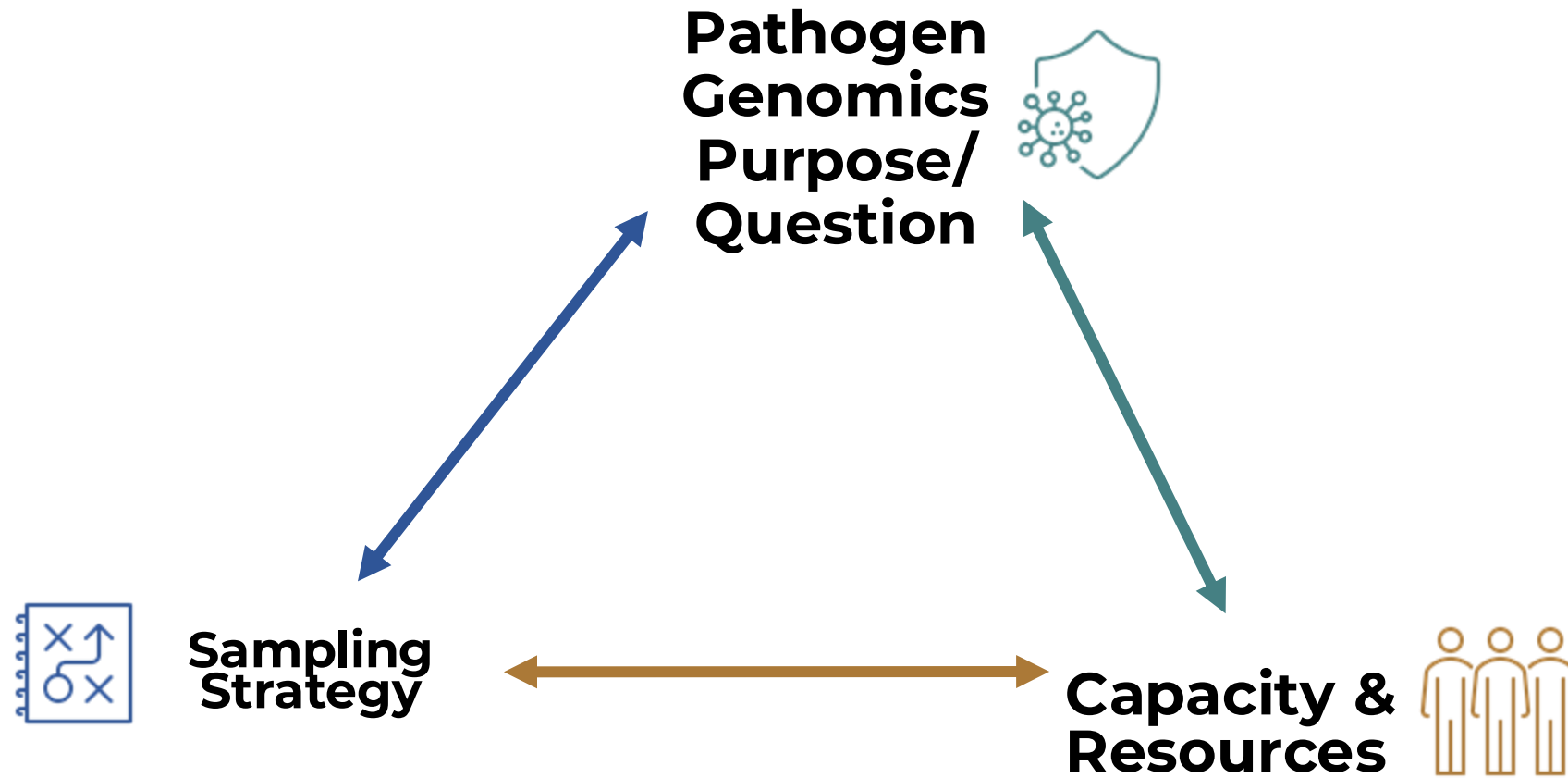
- Samples are collected based on their ease of accessibility rather than random selection; often used when resources or time are limited



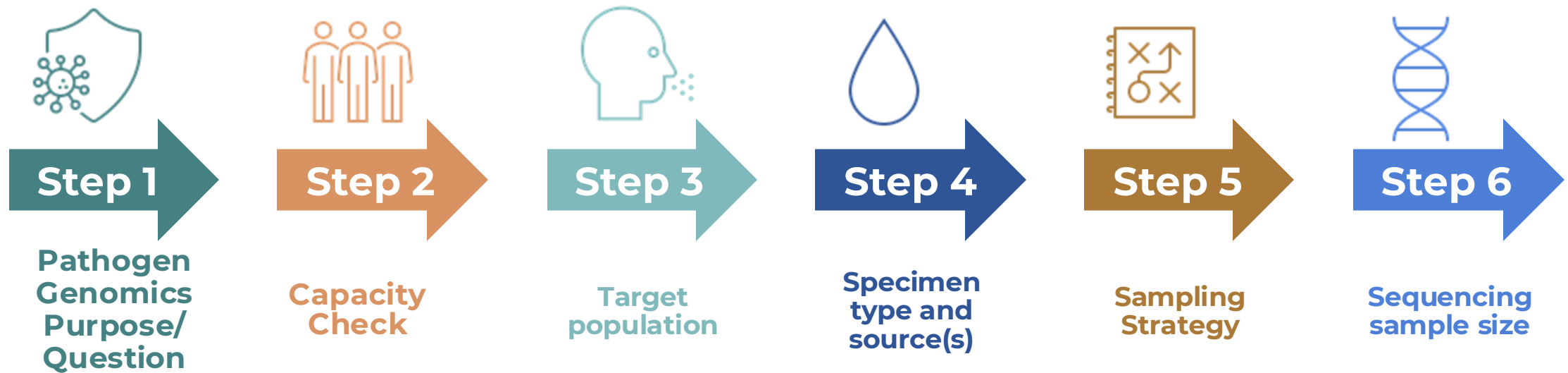
**Have you ever made a sampling strategy that you did not stick with? If so, why?**



# Genomic Surveillance Sampling Strategies



# Genomic Surveillance Sampling Steps





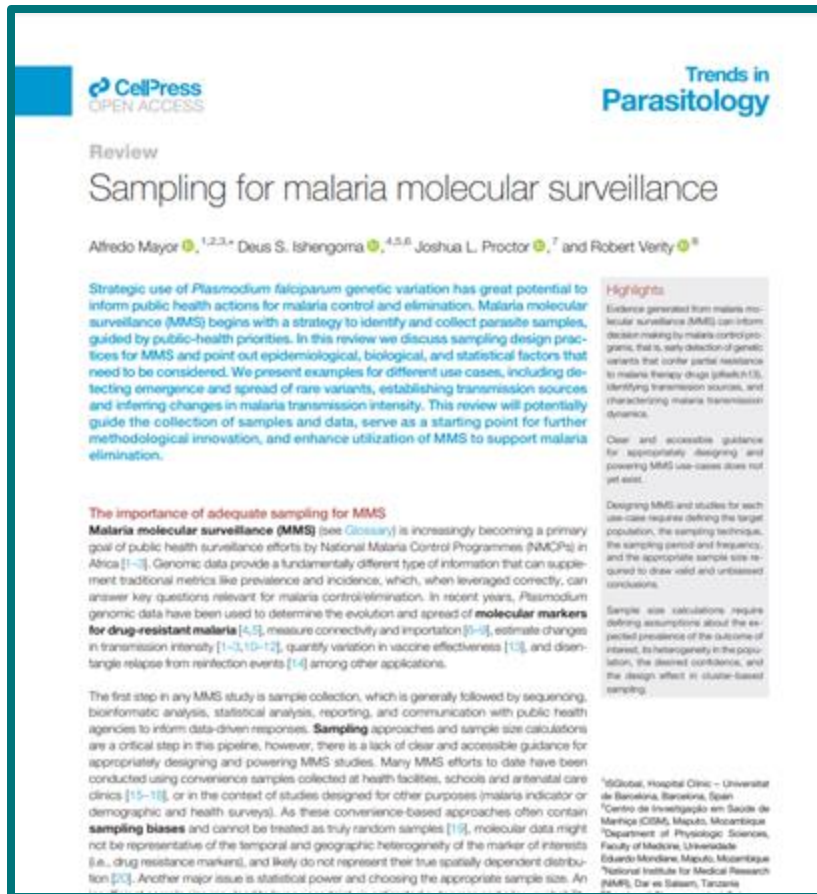
# Genomic Surveillance Sampling Strategies

Examples





# Case Study: Malaria



	1	2	3	4	5	6
	Use case	Goal	Target population	Sampling approach	Frequency	Parameters for sample size
	Early warning of biological threats	Detect emerging artemisinin resistance variants	Malaria clinical cases among children	Multi-cluster sampling	Cross-Sectional	1. Expected prevalence 2. Statistical power 3. Design effect
	Detection of transmission sources	Case classification (local/imported)	Whole population	Dense (all malaria cases in elimination settings)	Continuous	1. Pairs of infections 2. Genetic relatedness 3. Genetic resolution
	Inform transmission dynamics	Changes in transmission	Whole or sentinel population	All malaria cases in a representative area	Longitudinal	1. Genetic diversity 2. Infection clonality 3. Genetic resolution

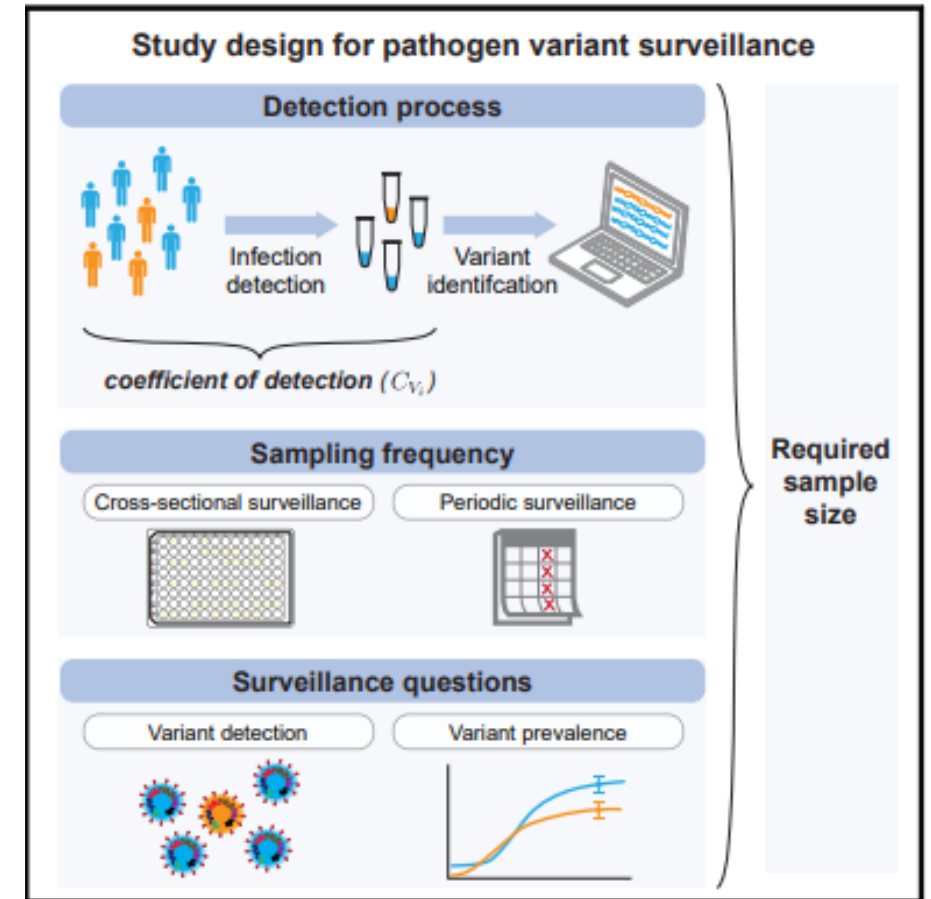
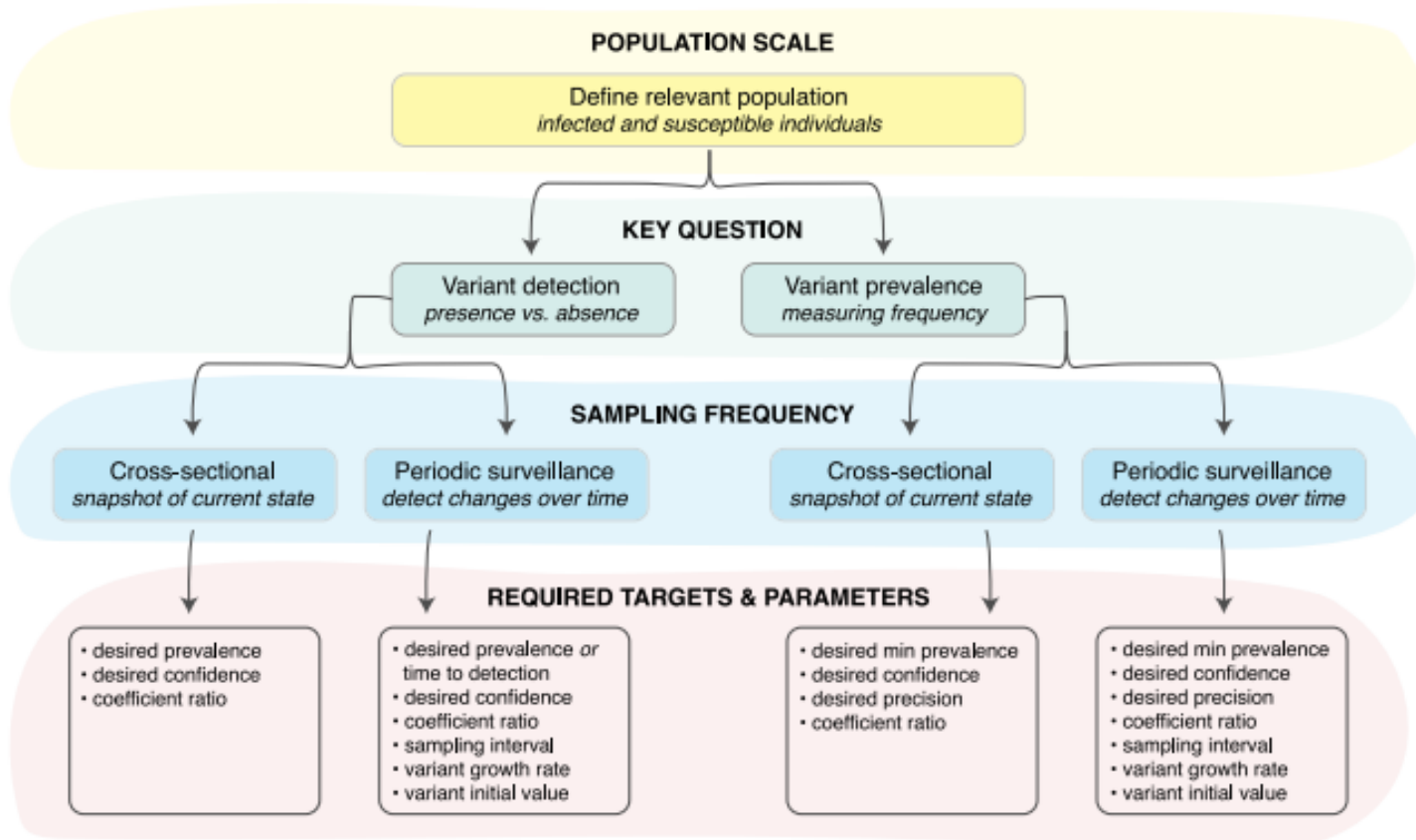
Trends in Parasitology

**Figure 1. Steps for designing a malaria molecular surveillance (MMS) approach.** To draw valid conclusions from MMS efforts, it is key to carefully decide how to select a sample that is representative of the target population. The surveillance purpose, and therefore the programmatic action expected from those efforts, will inform the relevant population to be sampled (which should be driven by the intervention target), the sampling method and the periodicity. All these parameters, which should be specific to the pre-defined population of interest as well as reflective of the logistical and biological sources of bias at the time of sampling, together with assumptions about the distribution of the marker of interest in the study population, need to be considered to calculate the appropriate sample size. Here we exemplify the different steps for three specific surveillance objectives: the detection of emerging variants of concern (such as mutations in *pfkdh13* associated with artemisinin resistance), the classification of cases as local or imported, and the detection of changes in transmission.

<https://www.sciencedirect.com/science/article/pii/S1471492223002118?via%3Dihub>



# Case Study: SARS-CoV-2



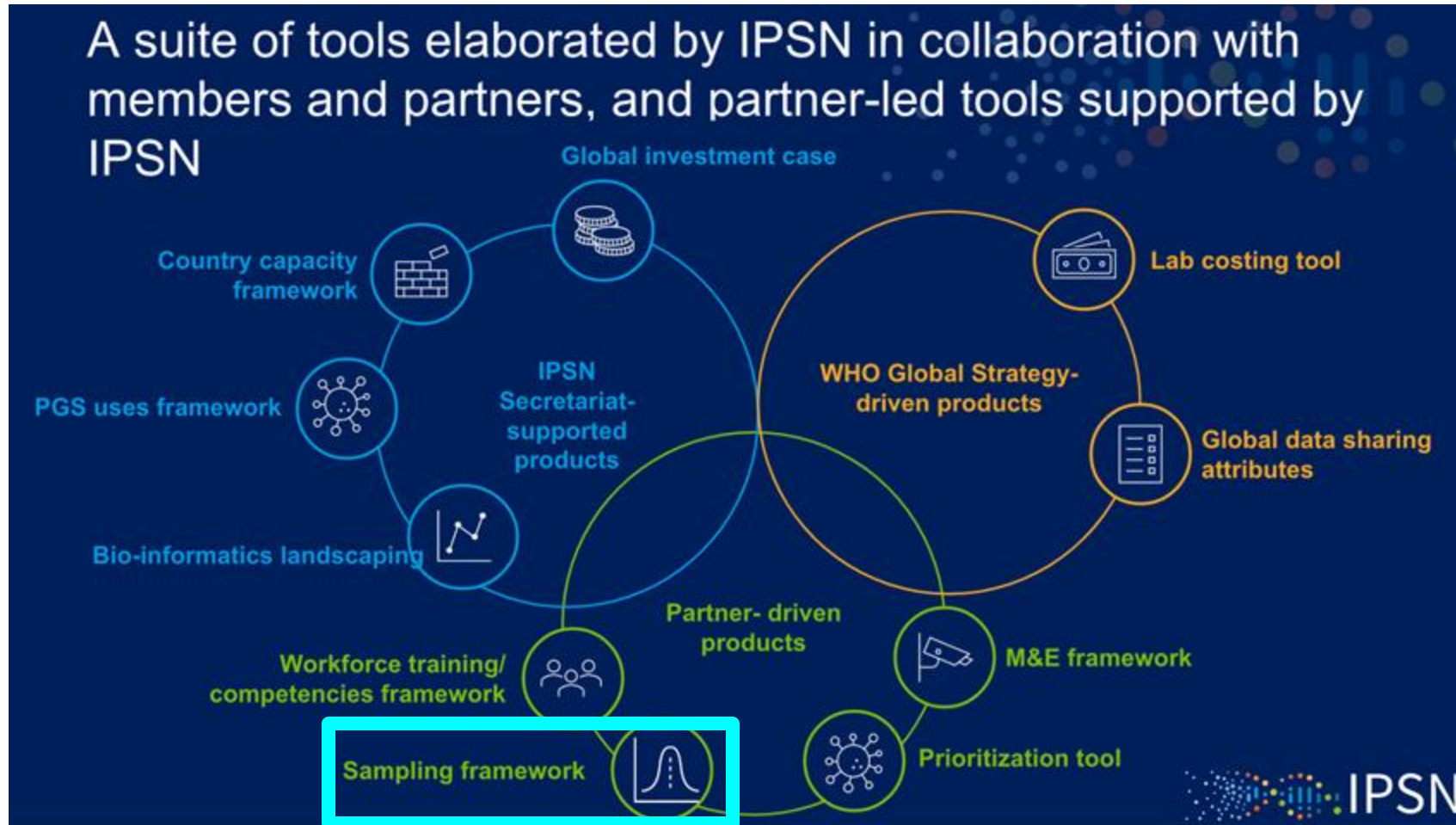


# Genomic Surveillance Sampling Strategies

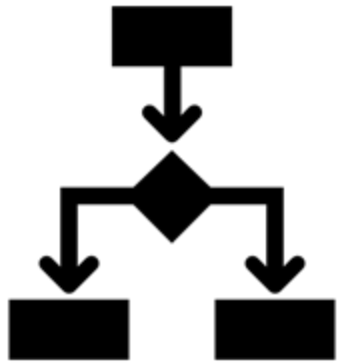
Partner Perspectives & Tool Building



# WHO IPSN global collaborative toolkit



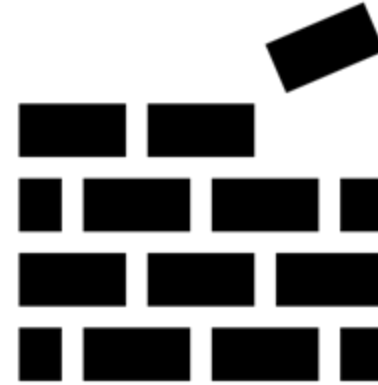
# Early Stage: Building Sampling Strategy Tools



**Decision  
tree**



**Sample size  
calculator**



**Building blocks  
& parameters**



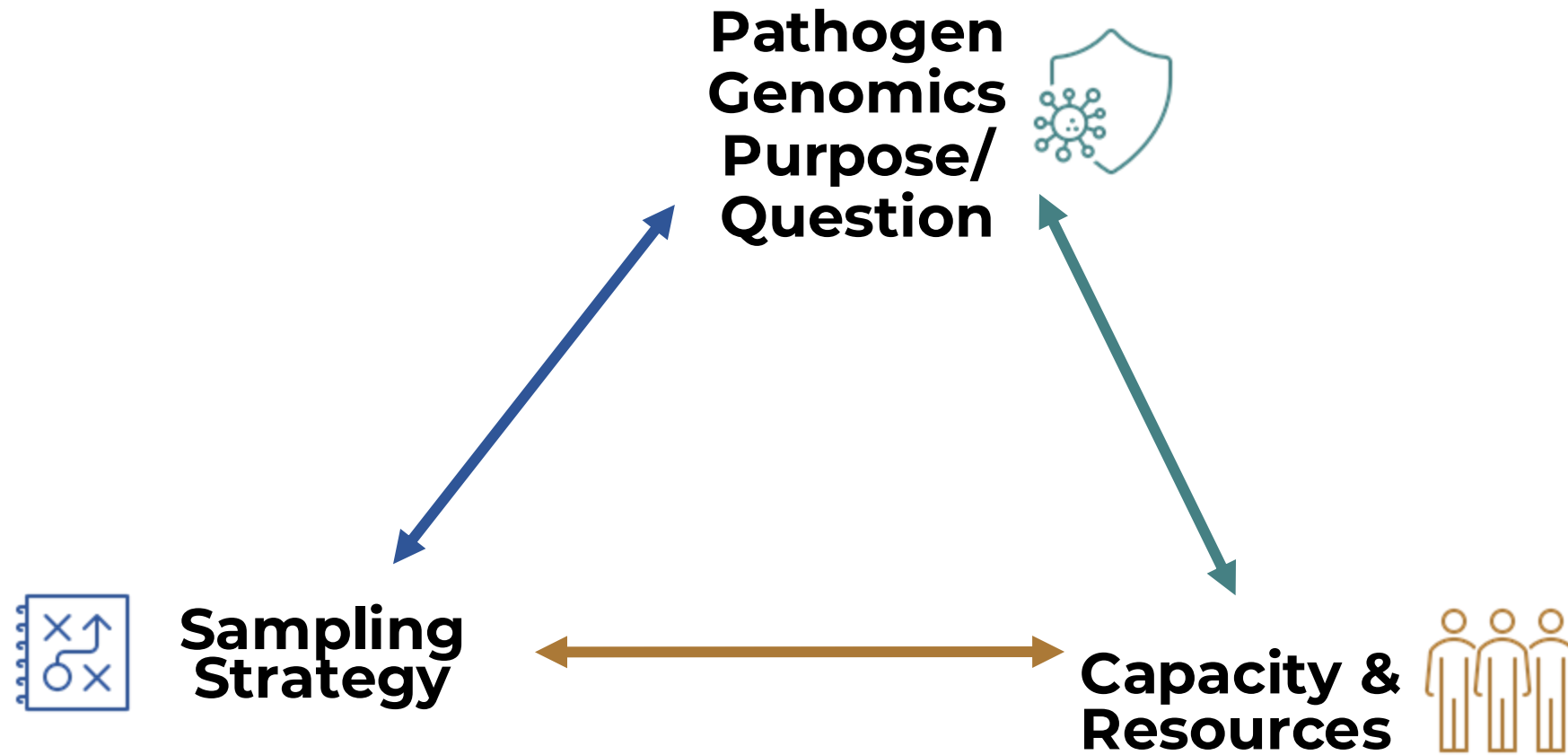
**Strategy  
resources**

## **framework**

- A tool or tools to support pathogen and capacity flexible genomic surveillance sampling strategies
- Actionable, easy-to-use, and sustainable tool or suite of resources



# Genomic Surveillance Sampling Strategies



# Genomic Surveillance Sampling Strategies

## **Pathogen Genomics Purpose/Question**



Early detection and warning systems

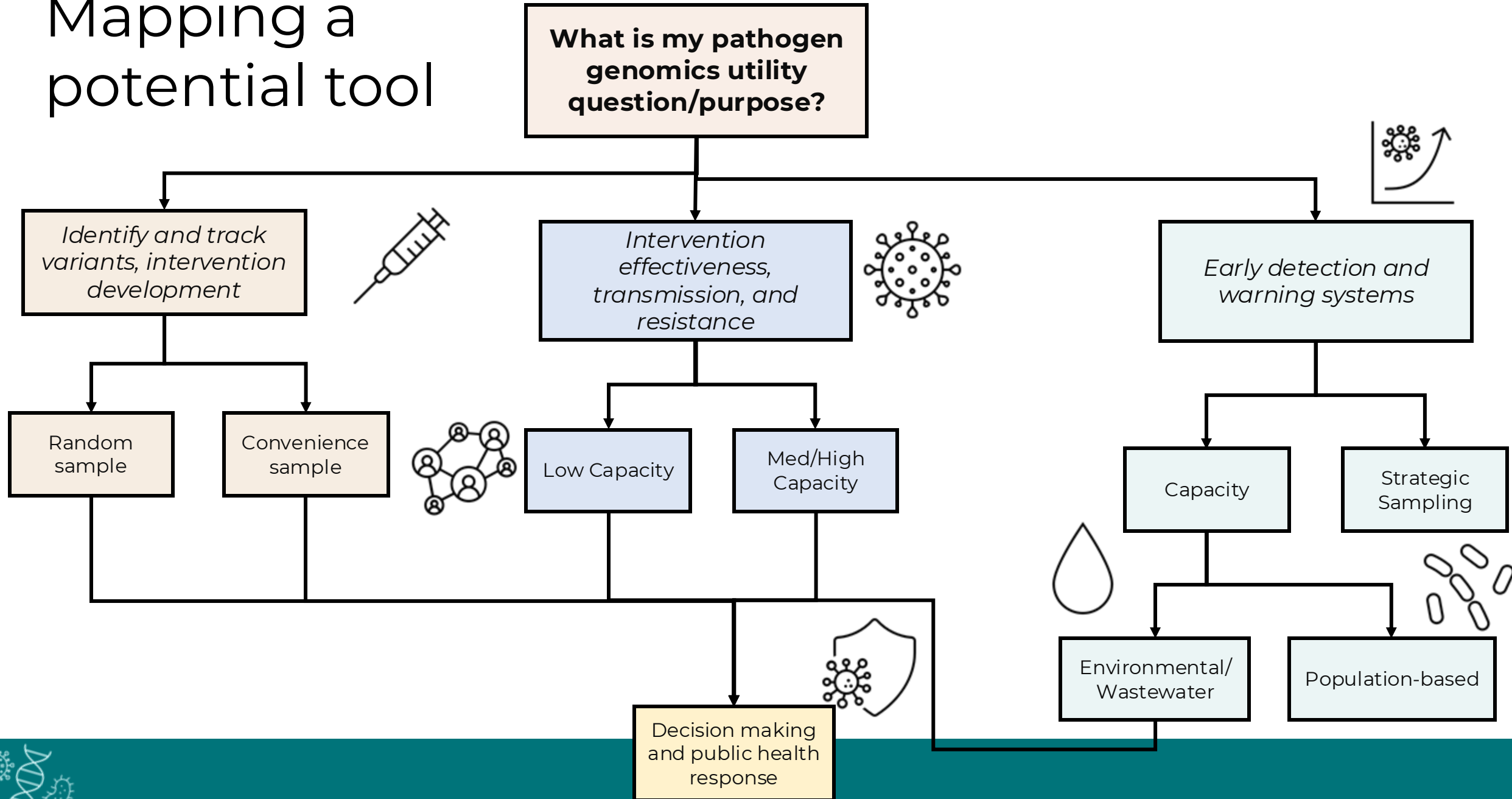
Identify and track variants, intervention development

Intervention effectiveness, transmission, and resistance



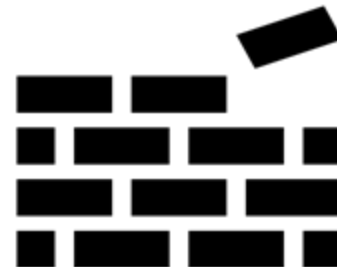
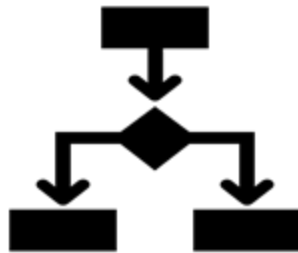
# Mapping a potential tool

## Decision Tree Tool



# Table Discussion:

1. What kinds of sampling strategy questions arise?
1. How is genomic sampling currently conducted in your country? Which stakeholders are part of these decisions (epi, lab, etc.)?
1. What kind of sampling strategy tools would be most helpful for your national planning?





# Sampling Strategy Tool Building

## Next Steps

1. Work with WHO IPSN and other partners to launch effort
2. Continue Global Technical Sampling Strategy WG
3. Develop sampling strategy tool(s)
4. Pilot tool(s) with partner countries

***Please reach out ([rix6@cdc.gov](mailto:rix6@cdc.gov)) if:***

- You'd like to join the Global Technical Sampling Strategy WG
- Interested in collaborating and piloting a future tool

