Leveraging the EMR: Tools & Rules

Presented by Bas de Veer

10:50pm-11:50pm

UW Husky Union Building

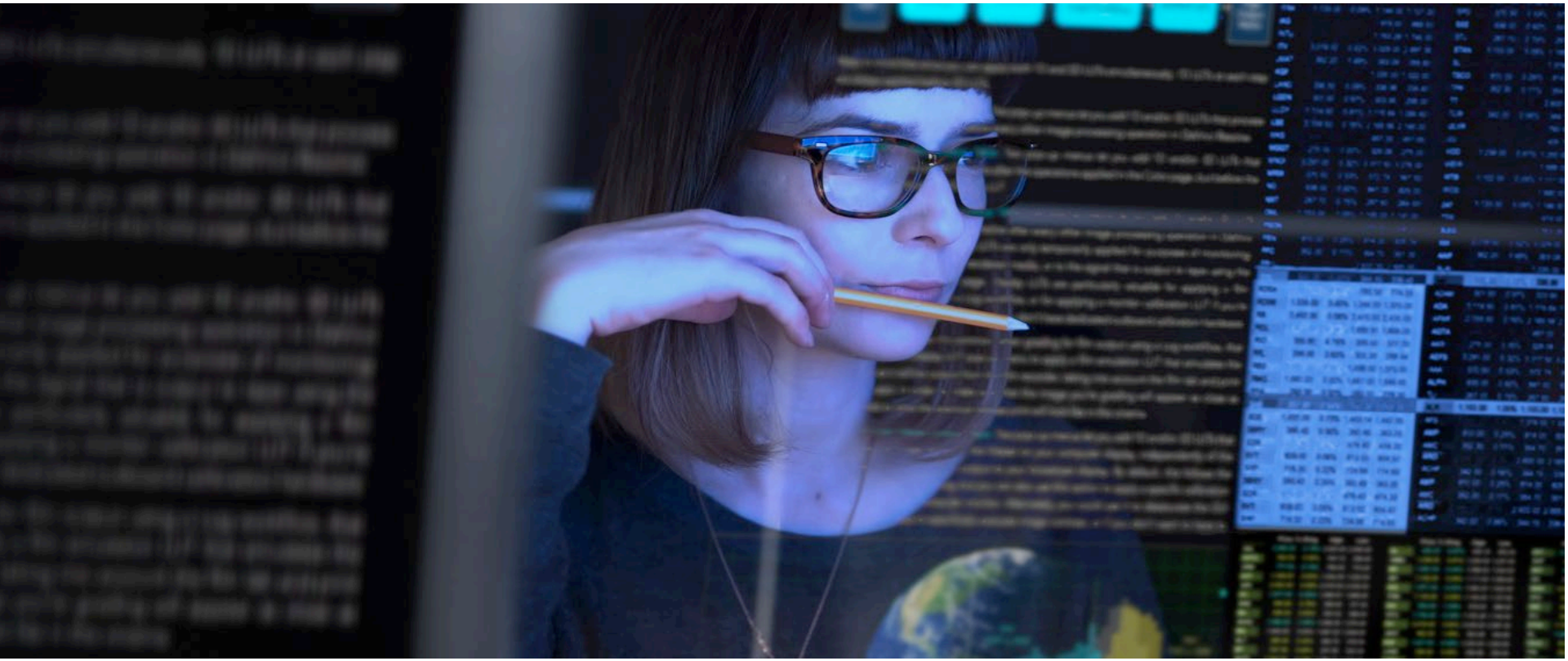
Room 145

ITHS | Institute of Translational Health Sciences
ACCELERATING RESEARCH. IMPROVING HEALTH.

UW MEDICINE | Analytics

Leveraging the EMR:

Tools & Rules for EMR Research Data Acquisition



Presented by: Bas de Veer, Bio-Medical Informatics Services Manager

UW Medicine

ITHS | Institute of Translational Health Sciences
ACCELERATING RESEARCH. IMPROVING HEALTH.

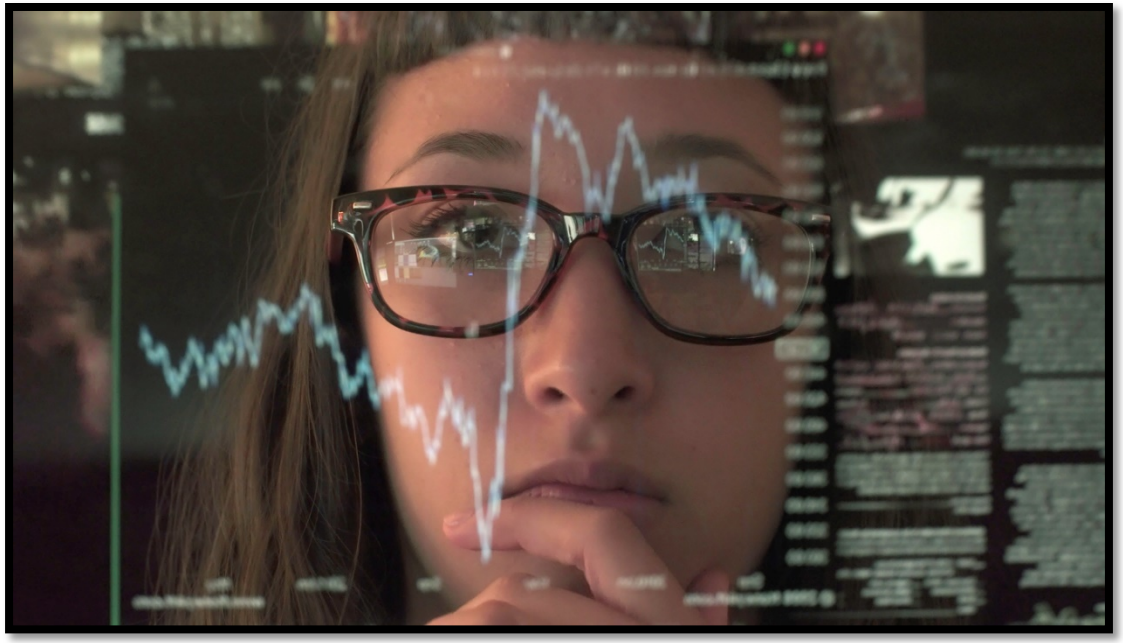
Learning Objectives

By the end of this session, you will be able to:

- Discuss the important information needed about your research question in order to effectively navigate the data access process
- Discuss the concepts of risks involved in the accessing of data
- Describe the various pathways or approaches to accessing different types of data
- Discuss the importance of data quality and validation, to be able to answer the question: Do you really have the cohort you need?

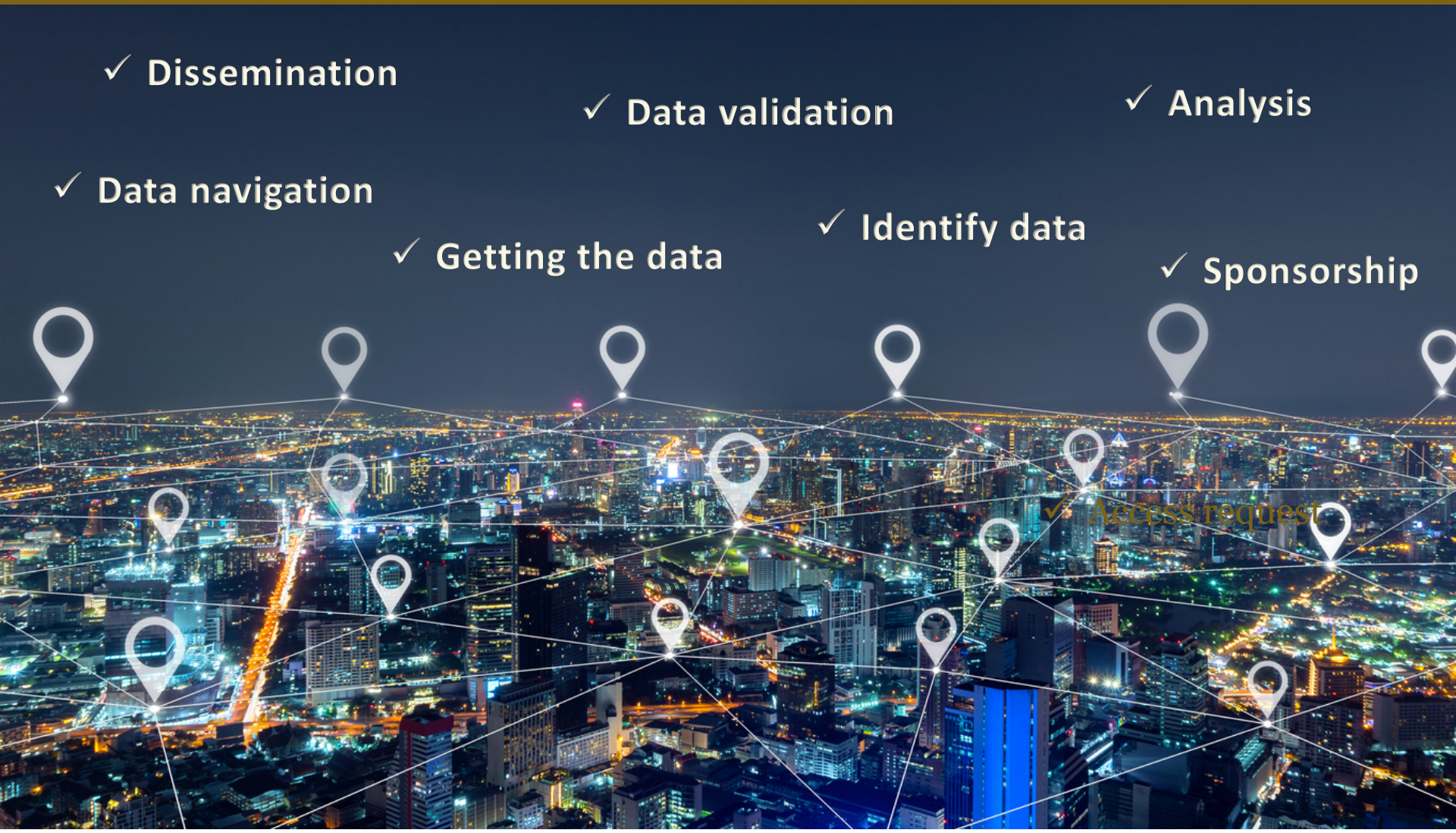
Assignment

Think about a research question you may have that will need health data…



(We'll come back to this in a few minutes)

How Do I Get Data to Support My Research?



UW Medicine

ITHS | Institute of Translational Health Sciences
ACCELERATING RESEARCH. IMPROVING HEALTH.

Worksheet

- Write down your research question and we will work through the areas together.

Identify Data

- What data do you need to answer your research question?
- Where does this data live?
- Is it structured or in notes? Data in notes is much harder to locate.
- Do you require Protected Health Information (PHI)? De-identified* patient data will be easier to obtain.
- A “limited data set” is another option if you absolutely must have dates.
- There are federal and state regulations that apply to health data, with substantial civil and criminal penalties for mishandling.

* Safe-Harbor method for de-identification requires removal of all of the following, without exception: Names, street address, city, zip (may keep first three digits of zip), all dates (may keep year), all ages > 89, telephone numbers, emails, any identifiers (SSN’s, MRN’s, IP addresses, health plan beneficiary numbers, etc.), Photos, fingerprints

Sponsorship

- This data is typically protected; ideally you will have a sponsor inside the organization.
- You may need to follow an IRB* process.
- You likely will need to follow an internal process as well.

NOTE: Obtaining an IRB does not grant you the right to access the data. The data is owned by the institution which has its own policies and processes for data release.

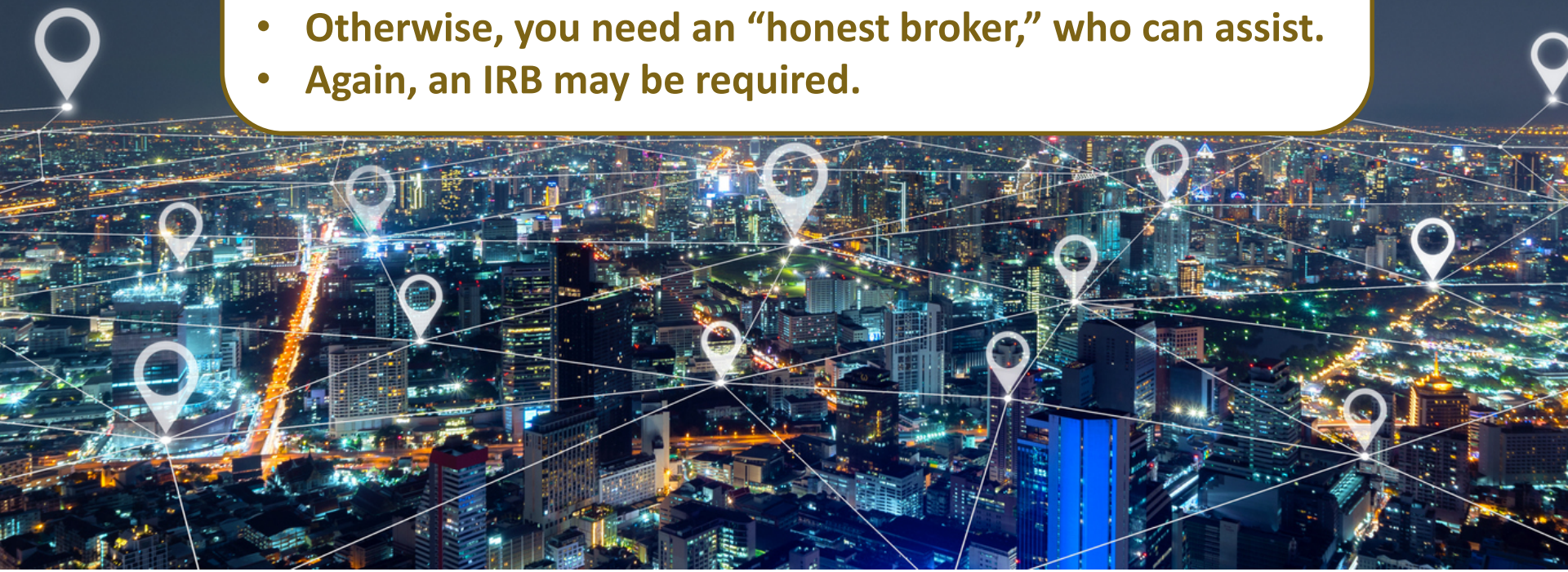
*IRB — Internal/Institutional Review Board. An IRB is a board, committee, or other group formally designated by an institution to review research involving humans as subjects. Typically, use of PHI for research requires human subject review.

Getting the Data

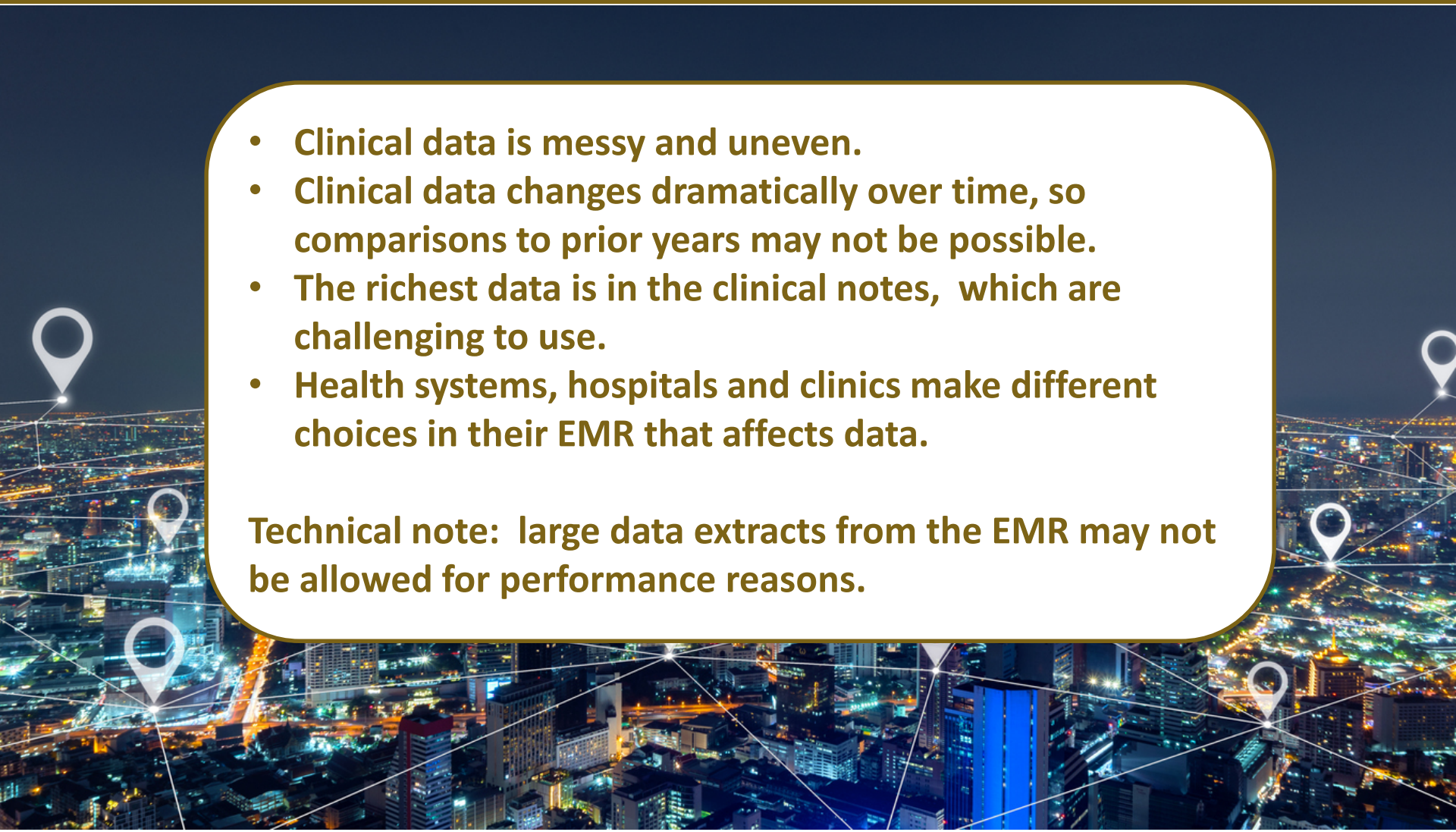
- If the data is in the EMR, you may need to write reports; if elsewhere, you may need SQL or other skills to extract.
- Locating the data you need is difficult (see data navigation) – you will likely need guidance on this from experts as well.
- You may also need training on EMR use, in order to understand the data entry process/scenario.
- At UW Medicine, we recommend partnering with an Affiliated Developer or Research IT.

Access Request

- You need to identify the system that has your data, and acquire access to that system.
- For direct access, you will need user login permission and access to the functions/data you've identified.
- Otherwise, you need an "honest broker," who can assist.
- Again, an IRB may be required.



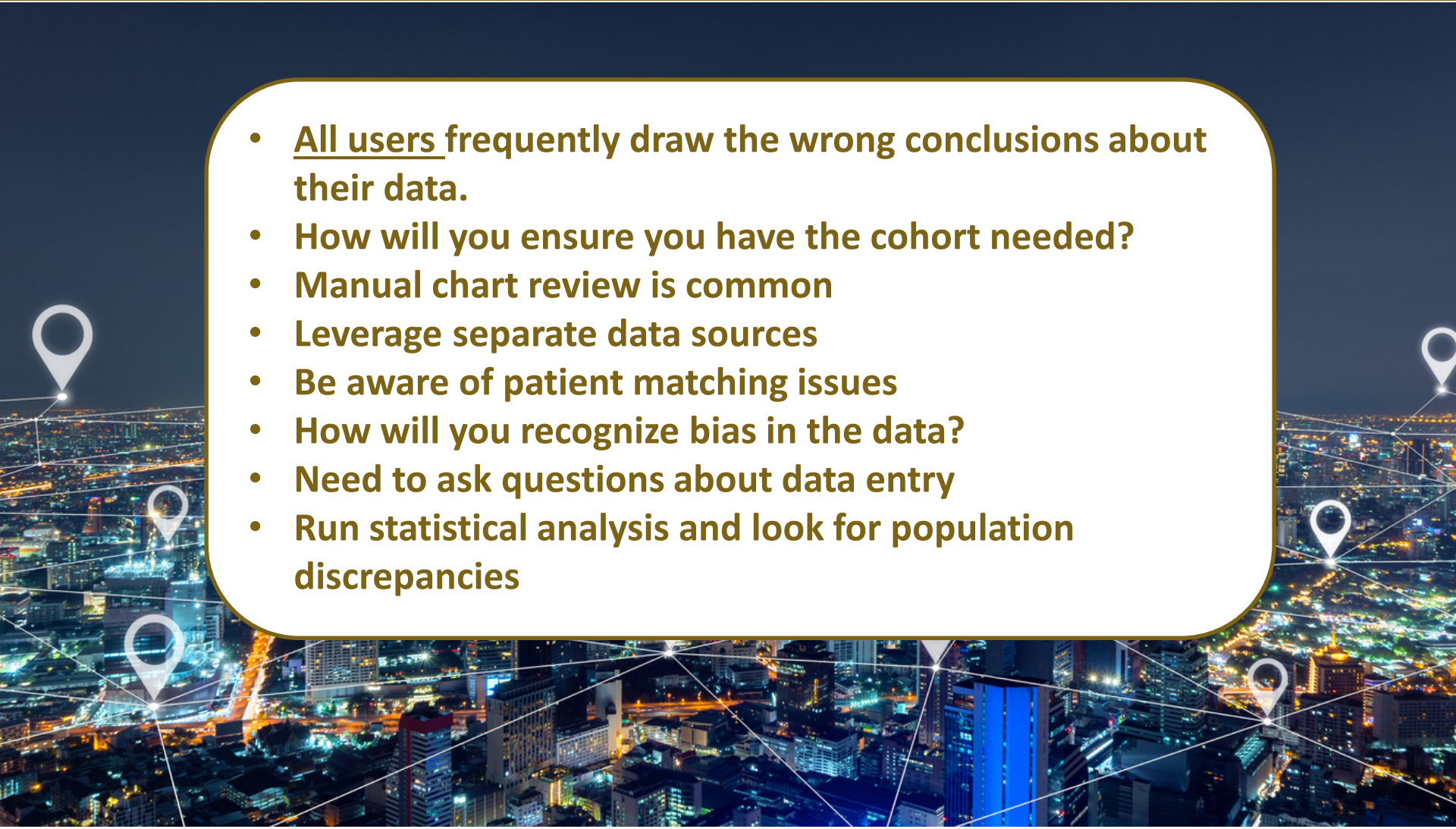
Data Navigation



- Clinical data is messy and uneven.
- Clinical data changes dramatically over time, so comparisons to prior years may not be possible.
- The richest data is in the clinical notes, which are challenging to use.
- Health systems, hospitals and clinics make different choices in their EMR that affects data.

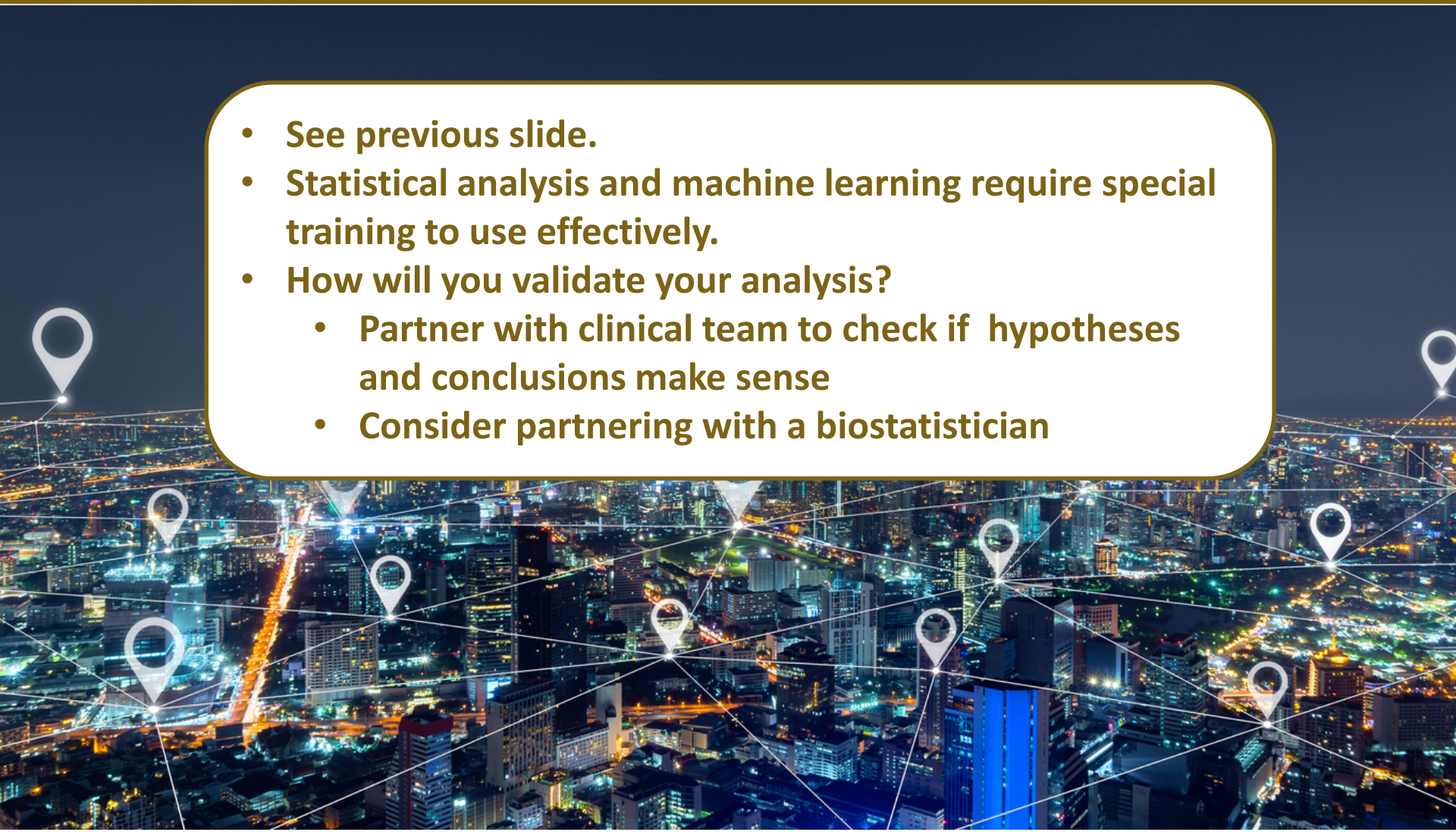
Technical note: large data extracts from the EMR may not be allowed for performance reasons.

Data Validation



- All users frequently draw the wrong conclusions about their data.
- How will you ensure you have the cohort needed?
- Manual chart review is common
- Leverage separate data sources
- Be aware of patient matching issues
- How will you recognize bias in the data?
- Need to ask questions about data entry
- Run statistical analysis and look for population discrepancies

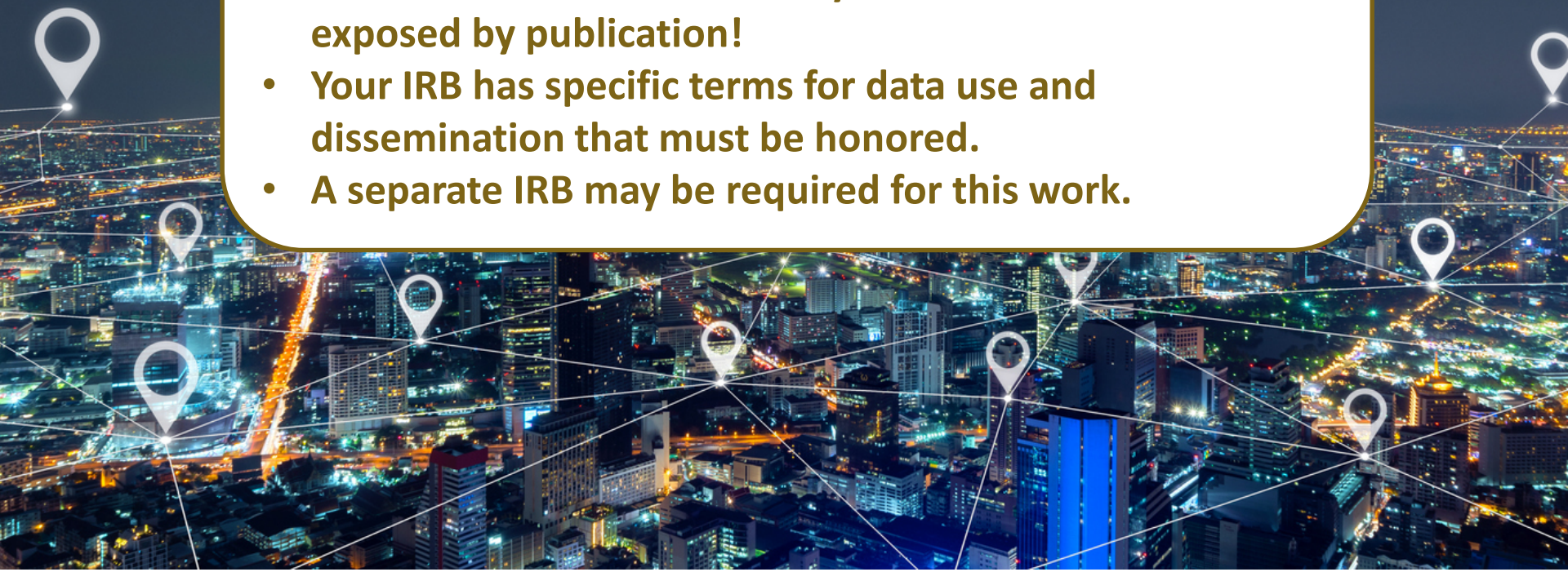
Analysis



- See previous slide.
- Statistical analysis and machine learning require special training to use effectively.
- How will you validate your analysis?
  - Partner with clinical team to check if hypotheses and conclusions make sense
  - Consider partnering with a biostatistician

Dissemination

- Data used for publication becomes part of public domain! You need to ensure this is permissible and may need to replicate research with publishable data.
- PHI and confidential health system data should never be exposed by publication!
- Your IRB has specific terms for data use and dissemination that must be honored.
- A separate IRB may be required for this work.



Learning Objectives

- Discuss the important information needed about your research question in order to effectively navigate the data access process
- Discuss the concepts of risks involved in the accessing of data
- Describe the various pathways or approaches to accessing different types of data
- Discuss the importance of data quality and validation, to be able to answer the question: Do you really have the cohort you need?