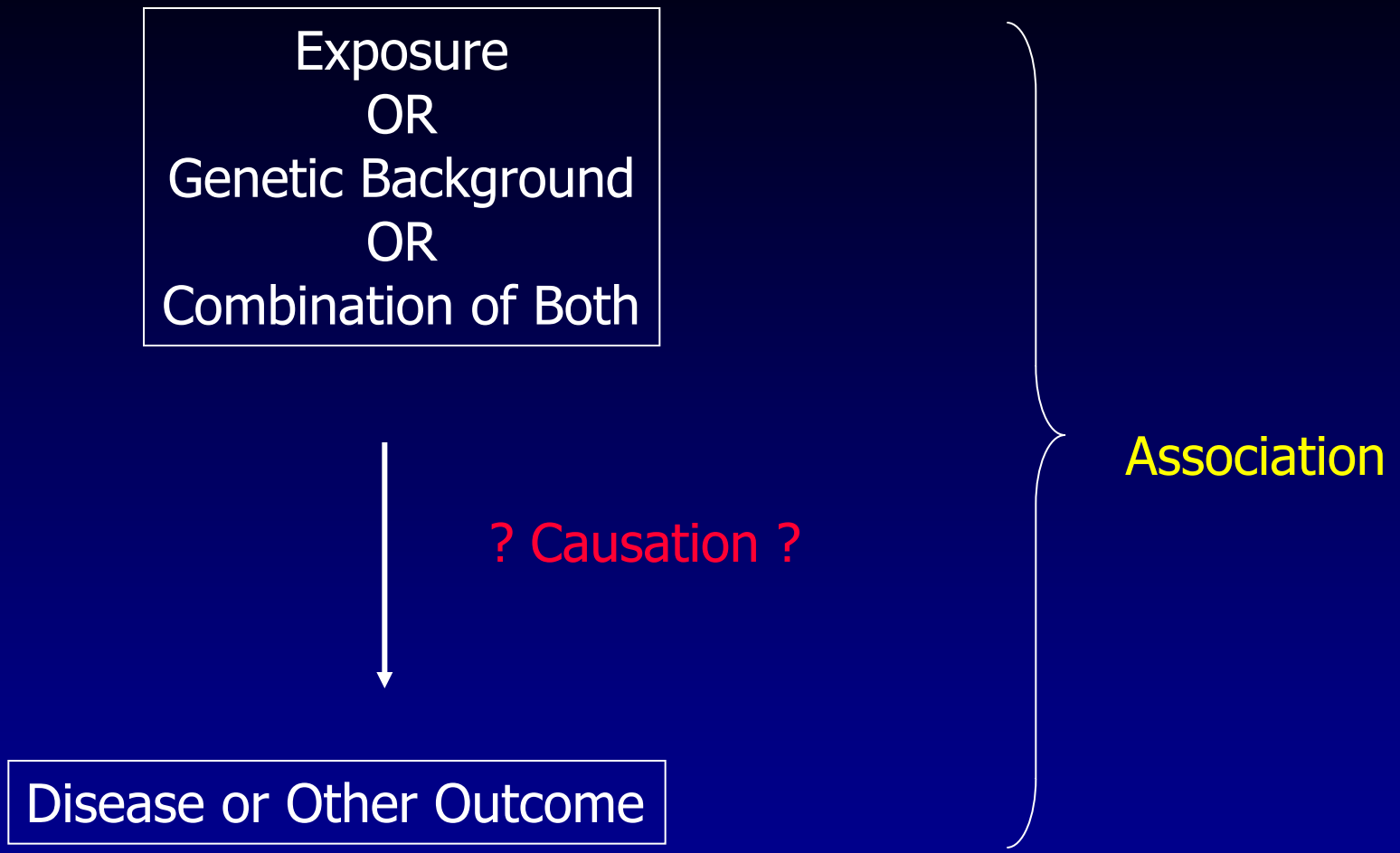
Association and Causation

Headlines

- Levels of causality
- Definitions
- Koch's postulates (1877)
- Hill's criteria (1965)
- Susser's criteria (1988, 1991)

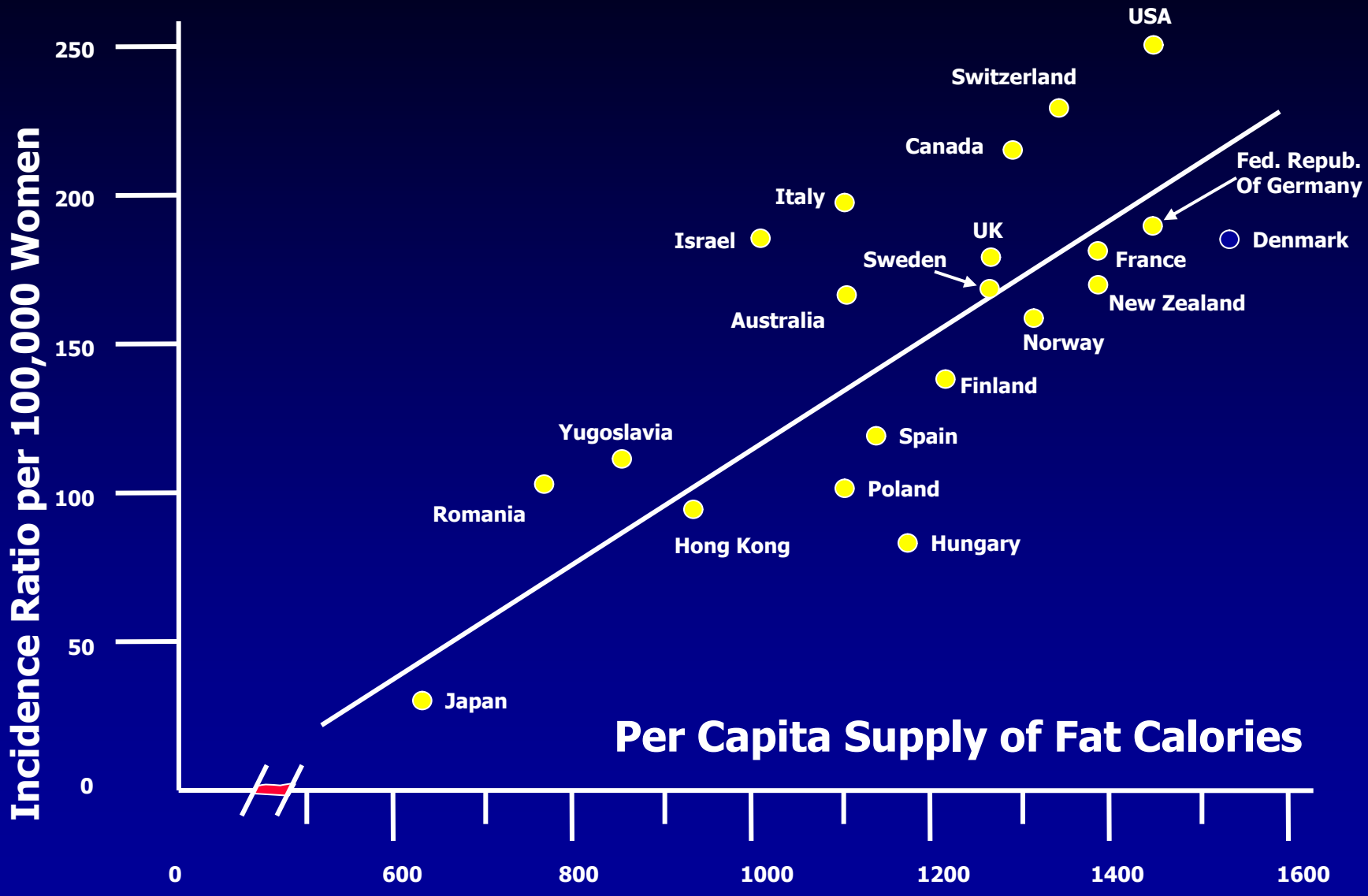
Relating

- **Exposures**: causes, risk factors
- **Outcomes:** effects, diseases, injuries, disabilities, deaths.
- **Statistical association** versus **biological causation:** cause-effect relationship



Suppose we determine that an exposure is associated with disease. How do we know if the observed association reflects a causal relationship?

Correlation between dietary fat intake and breast cancer by country.



Prentice RL, Kakar F, Hursting S, et al: Aspects of the rationale for the Women's Health Trial. J Natl Cancer Inst 80:802-814, 1988.)

Is there a relationship between breast cancer incidence and dietary fat consumption by country?

From the graph, we see that as average dietary fat consumption increases, breast cancer incidence increases.

What is wrong with this data?

The problem is the ecologic fallacy!

Prentice et al. J Natl Cancer Institute 1988 80:802-814

Why do an ecologic study? HYPOTHESIS Generation !

The data is easy to obtain, no follow-up or individual contact is needed.

An ecologic study can suggest way of research that may cast light on an etiologic relationship between exposure and disease.

HOWEVER,

an ecologic study does not itself demonstrate that a causal relationship exists.

Understanding Causality

- **Types of Association**
  - causal
  - noncausal
- **Types of Causal relationships**
  - direct
  - indirect
- **Types of causal factors**
  - sufficient
  - necessary

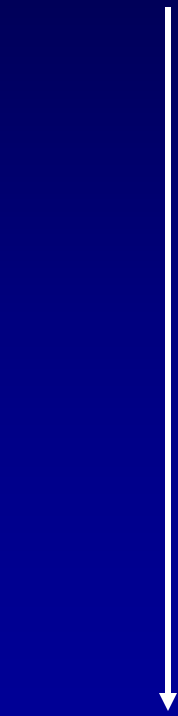
Two Types of Association: Real and Spurious

- A real association is present if
  the probability of occurrence of an event or
  the quantity of a variable depends upon the occurrence of one or more other events, characteristics or variables.
- Spurious associations refer to non-causal associations due to chance, bias, failure to control for extraneous variables (confounding), etc.

Interpreting Associations - Causal and Non-Causal

Causal

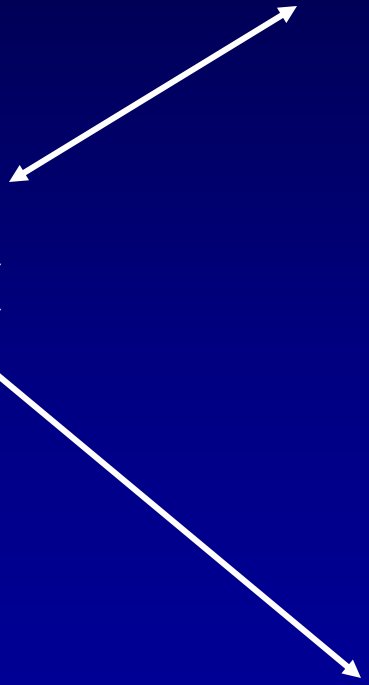
Characteristic Under Study



Disease

Non-Causal (due to confounding)

Characteristic Under Study



Factor X

Disease

The relationship between coffee consumption and pancreatic cancer

- In 1981, MacMahon et al. reported results from a case-control study of cancer of the pancreas.
- There was an apparent dose response relationship between coffee consumption and cancer of the pancreas, particularly in women.
- Was the disease caused by coffee consumption or by some factor closely related to coffee consumption?

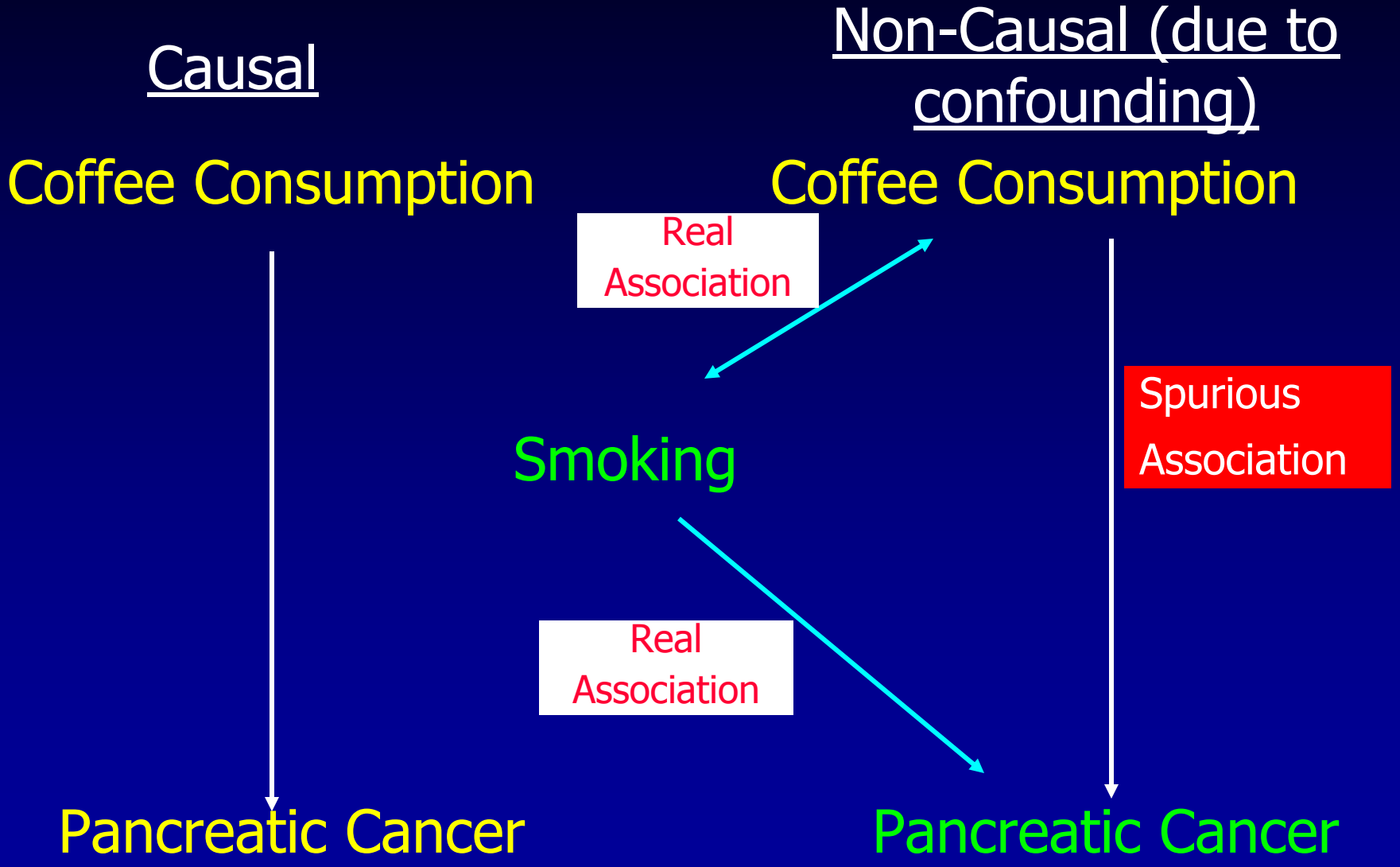
MacMahon B, et al. N Engl J Med 1981 304:630 - 33

The relationship between coffee consumption and pancreatic cancer

- Smoking is closely associated with both pancreatic cancer and coffee consumption.
- There were many issues with control selection and measurement of exposure levels in cases and controls.
- Subsequent studies were unable to reproduce the result.

MacMahon B, et al. N Engl J Med 1981 304:630 - 33

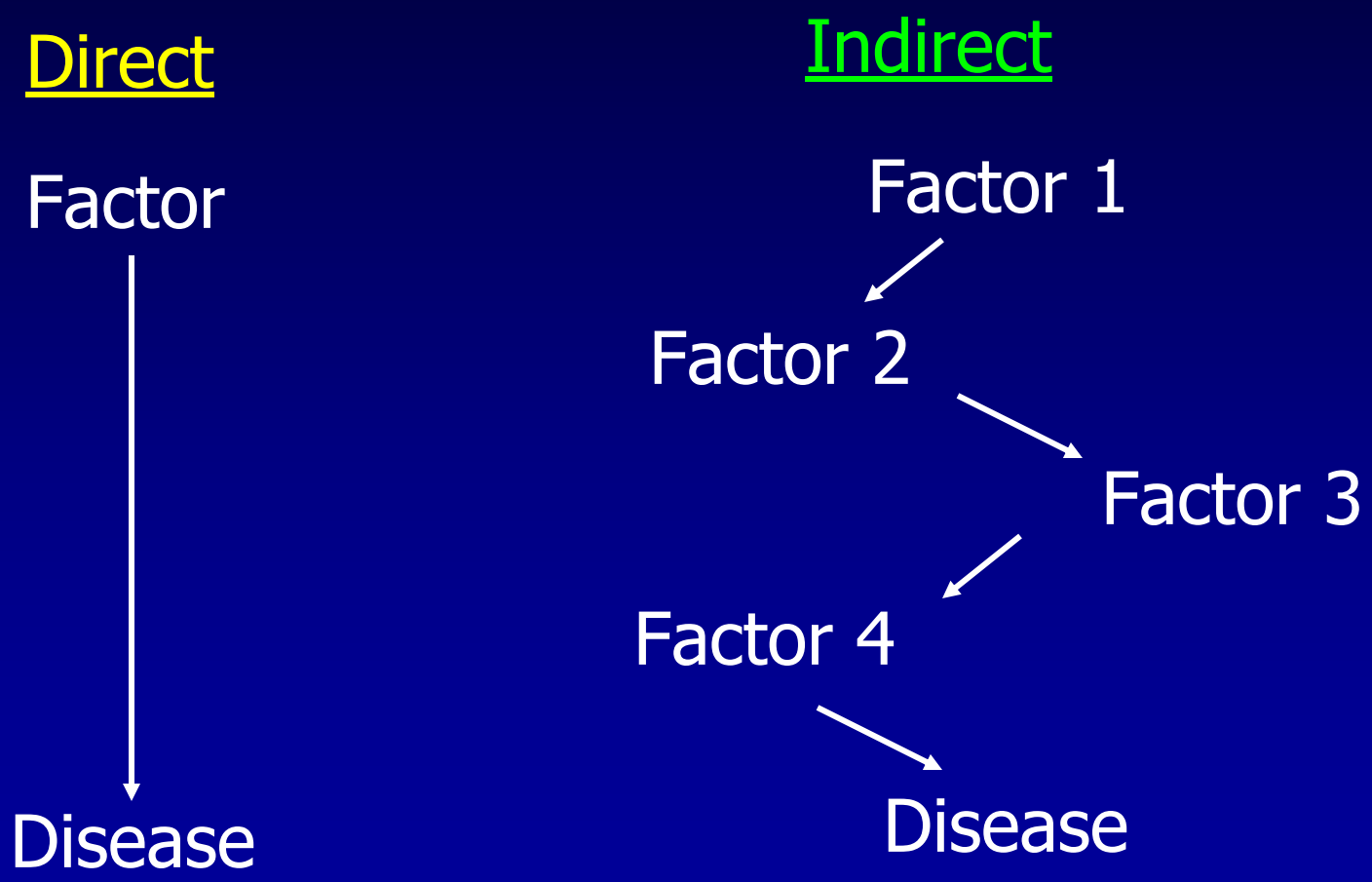
Interpreting Associations - Causal and Non-Causal



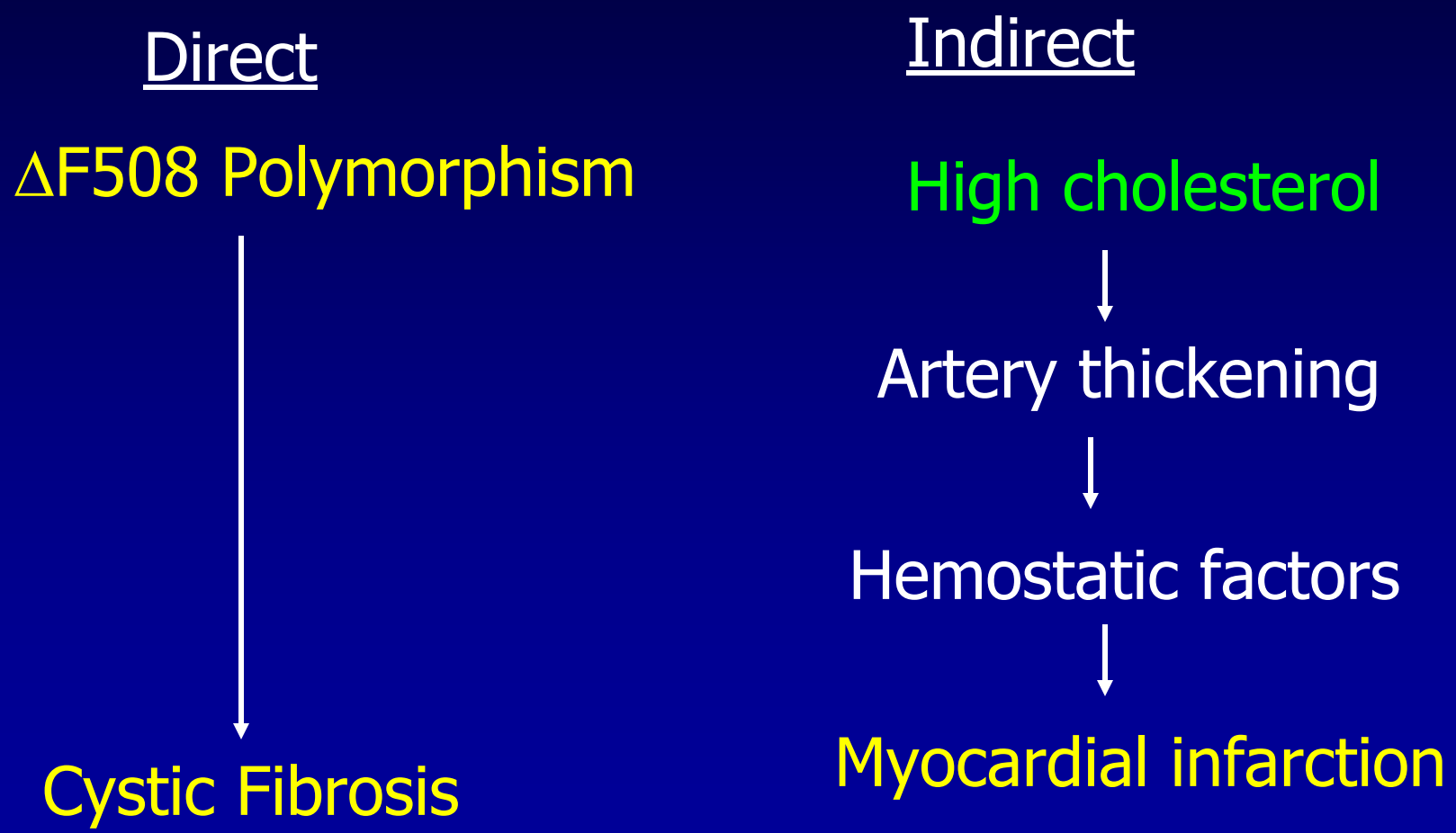
Why is it important to distinguish between causal and non-causal associations?

- Causal relationships are used to make public health decisions and design interventions.
- In our example, if smoking was causal, it would be irresponsible to target coffee drinking as an intervention.
- Very important to consider all confounders.

Types of Causal Relationships: Direct vs Indirect



Types of Causal Relationships: Direct vs Indirect



Four types of causal factors

- Necessary and sufficient
  - Without factor, disease does not develop
  - Example: HIV
- Necessary but not sufficient
  - Multiple factors, including main factor, required
  - Example: Development of tuberculosis requires M. tuberculosis and other factors, such as immunosuppression, to cause disease
  - Bacteria still necessary, but not sufficient to cause the disease

Four types of causal factors

- Sufficient but not necessary
  - Factor can produce disease, but not necessary
  - Example: Both radiation exposure and exposure to benzene are sufficient to cause leukemia, but neither are necessary if the other present.

- Neither sufficient nor necessary
  - Complex models of disease etiology
  - Example: High fat diet and heart disease, hypertension, diabetes, certain kinds of cancer

Understanding Causality

- you have determined:
  - there is a real association,
  - you believe it to be causal (ruled out confounding),
  - figured out that it is a direct causal factor
  - sorted out the necessary vs. sufficient factor issue…
- NOW have your proven CAUSALITY?

Henle-Koch's postulates (1877,1882)

Koch stated that four postulates should be met before a causal relationship can be accepted between a particular bacterial parasite (or disease agent) and the disease in question. These are:

1. The agent must be shown to be present in every case of the disease by isolation in pure culture.
2. The agent must not be found in cases of other disease.
3. Once isolated, the agent must be capable of reproducing the disease in experimental animals.
4. The agent must be recovered from the experimental disease produced.

Hill's Criteria (1897 - 1991)

The first complete statement of the epidemiologic criteria of a causality is attributed to Bradford Hill (1897 - 1991). They are:

1. Consistency (on replication)
2. Strength (of association)
3. Specificity
4. Dose response relationship
5. Temporal relationship (directionality)
6. Biological plausibility (evidence)
7. Coherence
8. Experiment

Consistency (I)

CONSISTENCY

Is the same association found in many studies? Hundreds of studies have shown that smoking and lung cancer are associated, and no serious study has failed to show this association. But whether oral contraceptives are associated with breast cancer is uncertain because some studies show an association, but others do not.

Consistency (II)

- Meta-analysis is an good method for testing consistency. It summarizes odds ratios from various studies, excludes bias
- Consistency could either mean:
  - Exact replication (as in lab sciences, impossible in epidemiological studies)
  - Replication under similar circumstances (possible)

Strength of Association

STRENGTH

Is the association strong? Heavy smoking is associated with a twenty-fold higher rate of lung cancer, and a doubled rate of coronary heart disease. The association of smoking with lung cancer is therefore stronger than its association with heart disease. The stronger the association the more likely it is to be truly causal.

Expressions of Strength of Association

- Quantitatively:
  - Effect measure (OR, RR): away from unity (the higher, the stronger the association)
  - P-value (at 95% confidence level): less than 0.05 (the smaller, the stronger the association)
- Qualitatively:
  - Accept alternative hypothesis: an association between the studied exposure and outcome exists
  - Reject null hypothesis: no association exists

Dose-response relationship (I)

Dose-response relationship

If a regular gradient of disease risk is found to parallel a gradient in exposure (e.g. light smokers get lung cancer at a rate intermediate between non-smokers and heavy smokers) the likelihood of a causal relationship is enhanced. **Dose-response** is generally thought of as a sub-category of **strength**.

Dose-response relationship (II)

Dose-response relationship

However, dose-response is not relevant to all exposure-disease relationships, because disease sometimes only occurs above a fixed threshold of exposure, and thus a dose-response relationship need not be seen. (remember also that misclassification of adjacent classes can easily produce an apparent dose-response relationship)

Time-order (temporality, directionality)

EXAMPLE 1.

Studies have found an inverse relationship between a person's blood pressure and a person's serum calcium. But which is the cause and which the effect?

Time-order can also be uncertain when disease has a long latent period, and when the exposure may also represent a long duration of effect.

Time order

EXAMPLE 2:

Low serum cholesterol has been linked to increased risk of colon cancer in prospective cohort studies. But is a low serum cholesterol a cause of colon cancer, or does an early phase of colon cancer cause low cholesterol levels?

Specificity of Outcome

EXAMPLE 1.

Asbestos causes a specific lung disease, asbestosis, distinguishable from many other lung diseases. But low level lead exposure is associated with lower IQ rather than a distinguishable brain syndrome. Thus lead is more uncertain as a cause because of possible confounding with other causes of this rather non-specific effect, low IQ (e.g. SES).

Coherence

- **Theoretical:** compatible with pre-existing theory
- **Factual:** compatible with pre-existing knowledge
- **Biological:** compatible with current biological knowledge from other species or other levels of organization
- **Statistical**: compatible with a reasonable statistical model (e.g. dose-response)

Biological Coherence (I)

EXAMPLE:

Presence of a serological marker of hepatitis B infection is associated (in Asia at least) with greatly elevated rates of liver cancer. That Hepatitis B infection is a true cause of liver cancer is also supported by the finding of the viral genome in many liver cancers.

Biological Coherence (II)

By contrast, Reserpine (an anti-hypertensive drug) was thought to be a cause of breast cancer based on some studies done in the early 1970's. But there was no other supporting biological information, or any truly plausible biological mechanism. Subsequent larger studies failed to support this association. Similarly for EMF and carcinogenesis.

Susser's criteria (I)

- Mervyn Susser (1988) used similar criteria to judge causal relationships.
- In agreement with previous authors, he mentioned that two criteria have to be present for any association that has a claim to be causal: i.e. time order (X precedes Y); and direction (X leads to Y).

Susser's Criteria (II)

- Rejection of a hypothesis can accomplished with confidence by only three criteria:
  - time order
  - Consistency
  - factual incompatibility or incoherence.
- Acceptance can be achieved by only four:
  - Strength
  - Consistency
  - predictive performance
  - statistical coherence in the form of regular exposure/effect relation.

CAUSAL CRITERIA COMPARED

SURGEON GENERAL	SUSSER	BRADFORD-HILL
	ASSOCIATION	DOSE RESPONSE*
	DIRECTION	EXPERIMENT
TIME ORDER	TIME ORDER	TIME ORDER**
STRENGTH	STRENGTH	STRENGTH
CONSISTENCY	CONSISTENCY	CONSISTENCY
SPECIFICITY	SPECIFICITY	SPECIFICITY
COHERENCE	COHERENCE	COHERENCE***
	PREDICTIVE PERFORMANCE	

*Included under strength in other criteria. ** Temporality in Bradford-Hill. *** Biological plausibility in Bradford-Hill

Asbetos and Lung Cancer

Well - established temporal relationship

Asbestos

Latent period of 10 - 20 yrs

Lung Cancer

New Study

Asbestos

Latent period of 3 yrs

Lung Cancer

In this case, the latent period is not long enough for lung cancer to develop if caused by exposure.

Associations are observed
Causation is inferred

It is important to remember that these criteria provide evidence for causal relationships.

All of the evidence must be considered and the criteria weighed against each other to infer the causal relationship.

Thank You