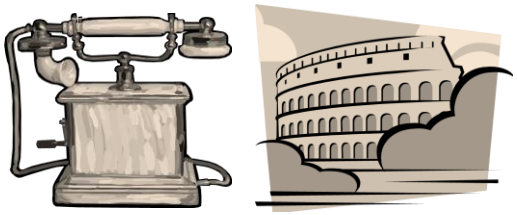
AMPHI
Academic Public Health Initiative

UMC St Radboud

# Tele-ARENA

(**Tele**fonisch **A**angenaam **R**efereren **E**n **N**etwerken **A**MPHI)



## 2. Onderzoeksmethode

*van doelstelling naar dataverzameling*


Alma Tostmann

Epidemioloog @ Academische Werkplaats AMPHI

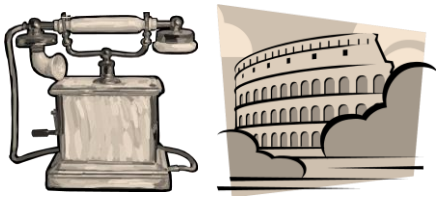
Afdeling Eerstelijnsgeneeskunde, UMC St Radboud, Nijmegen


*11 juni 2013*


Acknowledgement:

Presentations from EPIET Intro Course 2012 | From study objectives to analysis plan & From analysis plan to data collection – H. Maguire

# Welkom bij de 2e Tele-ARENA



- Tele-ARENA = telefonische nascholing opzet wetenschappelijk onderzoek
- Onderwerpen + praktijkvoorbeeld:
  1. **Het onderzoeksvoorstel**: van onderzoeksvraag naar doelstellingen
  2. **Dataverzameling**: van doelstellingen naar dataverzamelingsplan
  3. **Data analyse**: beschrijvende & statistische analyse – een opfrisser
  4. **Onderzoeksrapportage**: data presentatie, kernboodschap, aanbevelingen

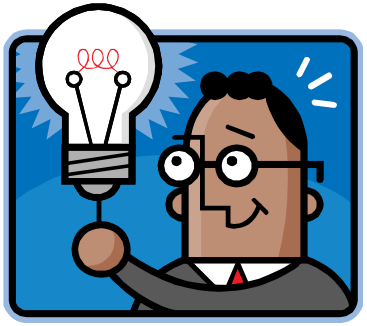
**Tijdens de presentatie:**

- Start: deelnemers zetten telefoon op **luisterstand: intoetsen "16"**
- Vragen tussendoor mag (eerst telefoon uit luisterstand zetten: 16)

# Wrap-up 1e Tele-ARENA

- Studieprotocol geeft onderzoeker richting:
  - **voor** (plannen), **tijdens** (data verzamelen, analyseren) en **na** (conclusies, aanbevelingen) het onderzoek
- Introductie schrijven met Trechtermethode (algemeen → specifiek)
- **Centraal in het hele onderzoeksproces**
  - Doel: overkoepelend doel van de hele studie
  - Vraagstellingen – waar je antwoord op wilt krijgen dmv het onderzoek:
    - Geven focus aan het onderzoek
    - Zijn bepalend voor de onderzoeksmethode
    - Voorkómen verzamelen van data die niet nodig is
    - Geven richting aan data analyse, interpretatie en aanbevelingen
- Goed geformuleerde vraagstellingen
  - Zijn SMART: **S**pecifiek **M**eetbaar ***A****chievable* (haalbaar) **R**ealistisch **T**ijdsgebonden
  - En geven aan **<u>wat</u>** je wilt meten, in **<u>wie, waar</u>** (*specifiek, meetbaar*), **<u>hoe</u>** (*haalbaar, realistisch*) en **<u>wanneer</u>** (*tijdsgebonden*)

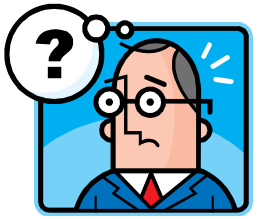
# De vraagstelling is er, en dan....



**<u>Zorg dat de onderzoeksvraag kraakhelder is</u>**

1. Welke onderzoeksvorm
2. Centrale variabelen en indicator
3. Selectie onderzoekspopulatie
4. Analyseplan
5. Benodigde data elementen
6. Methode data verzameling
7. Uitwerking dataverzameling
8. Dataverzamelingsplan

Indien nodig
**vraagstelling bijwerken:**
als dataverzameling niet aansluit bij vraagstelling is studie zinloos



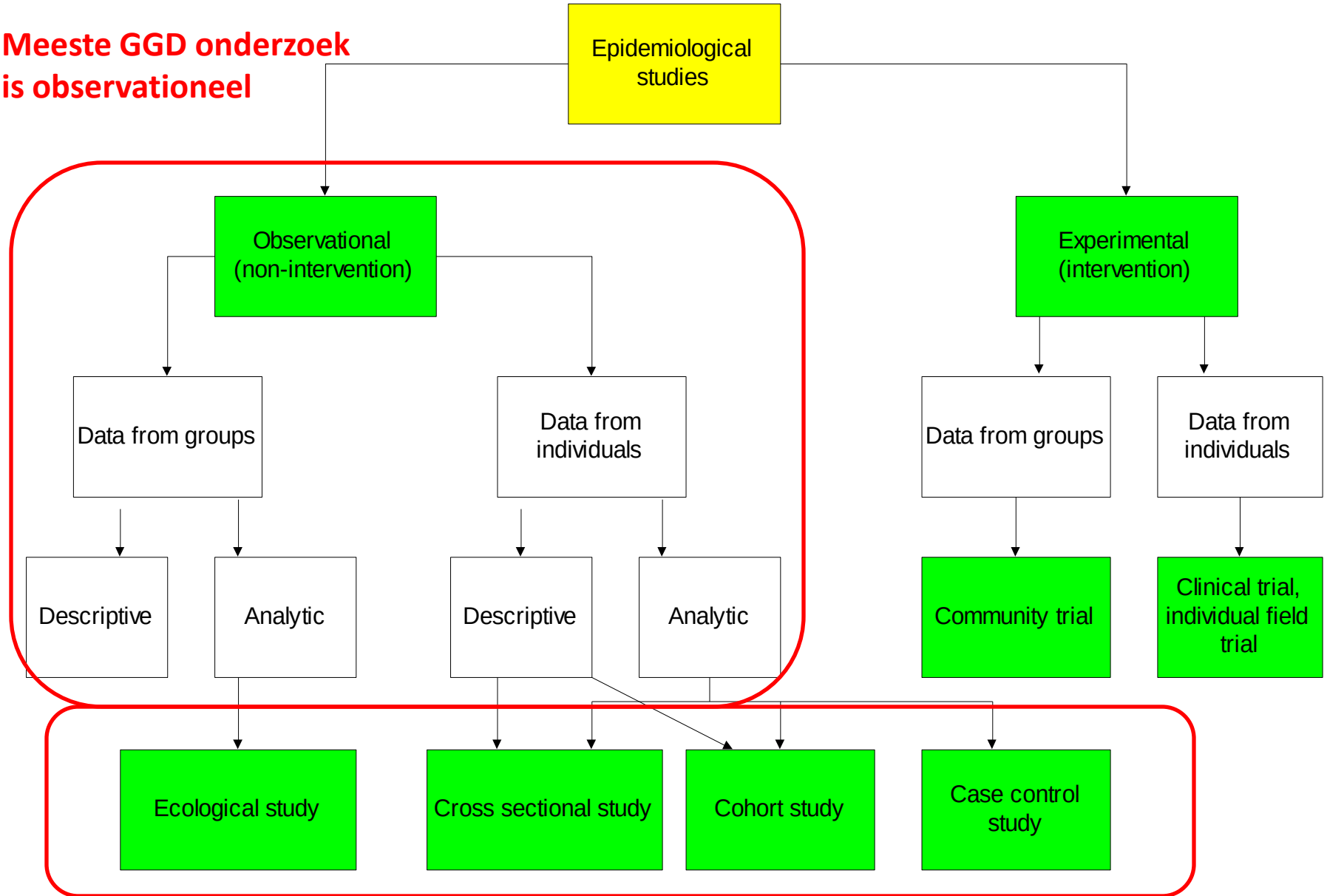
# 1. Onderzoeksvorm

Keuze onderzoeksvorm hangt af van o.a. volgende zaken:

- **Observationeel / experimenteel**
  - GGD onderzoek: meestal observationeel, dus geen medische interventie
  - Experimenteel = interventie
    - Hepatitis B vaccinatie bij non-responders van Stijn Raven
- **Data van groepen of individuen**
  - Individuen: personen - ziek ja/nee, gevaccineerd ja/nee,
  - Groepen: kinderdagverblijven die wel/niet een nieuwe Toolkit hebben gehad
- **Beschrijvend of analytisch**
  - Beschrijvend: schatten van een percentage
    - Percentage hepatitis B/C dragerschap bij immigranten die voor TBC worden gescreend (J Klomp)
  - Analytisch: vergelijken van groepen
    - Motiverende/belemmerende factoren GGD personeel t.a.v. opname Varicella vaccinatie in RVP (H Ruijs)
- **Is de ziekte acuut of chronisch**
  - Prevalentie schatten voor chronische ziekten (bv. HIV seropositiviteit)
  - Incidentie data voor acute ziekten
- **Hoe vaak komt ziekte voor?**
  - Zeldzaam: case-controle studie
  - Veel voorkomend: cohort of cross sectionale studie

# 1. Onderzoeksvorm

**Meeste GGD onderzoek is observationeel**



# Voorbeelden onderzoeksvorm

- Risicofactoren voor slechte *compliance* malariaprofylaxe bij reizigers die malariaprofylaxe geadviseerd hebben gekregen door GGD Nijmegen?
  - Observationeel; individueel; analytisch; niet zeldzaam

  → **Cohort studie** mogelijke risicofactoren vergelijken tussen reiziger die wel versus niet hun profylaxe hebben gebruikt

- Verhoogt de *Vlekjes-ziekte Toolkit* het percentage kinderdagverblijven dat melding maakt van vlekjesziekten in regio Arnhem?
  - Experimenteel; groepen

    → **Community trial** - Vergelijken percentage KDVs dat melding maakt van vlekjesziekten in Toolkit groep vs. non-toolkit groep

- Wat is de attitude van GGD-IZB-medewerkers en CB-medewerkers t.a.v. opname van Varicella vaccinatie in het RVP en door welke factoren wordt deze attitude beïnvloed?
  - Observationeel, individueel, beschrijvend/analytisch

    → **Cross-sectionele studie**: beschrijven attitude, en vergelijken 'risico'factoren tussen positieve versus negatieve Varicella-RVP attitude

# 2. Centrale variabelen en indicator

**Indicator:**

- De **primaire uitkomstmaat** van je onderzoek
- Ofwel, het getal dat centraal staat in het antwoord op de onderzoeksvraag
- Gebaseerd op de **centrale variabelen**, meestal:
  - Ziekte: ja/nee
  - Risicofactor: ja/nee

Voorbeelden **indicatoren**

- Beschrijvend: absolute maat
  - Percentage (vaccinatiegraad op antroposofische school)
  - Gemiddelde: gemiddelde leeftijd van bezoekers SOA poli
- Analytisch: relatieve maat
  - RR of OR: associatie risicofactoren en ziekte
  - Verschil: gemiddelde leeftijd SOA-poli bezoekers met chlamydia versus gonorroe

Voorbeelden **centrale variabelen** – nodig om indicator te berekenen: bv ziekte, risicofactor, leeftijd

- Vaccinatiegraad school: vaccinatiestatus per kind, totaal aantal kinderen
- Gemiddelde: leeftijd van SOA poli bezoekers.
- RR/OR: ziek ja/nee; risicofactor ja/nee
- Verschil: leeftijd van SOA polibezoekers; testuitslag chlamydia/gonorroe

- Bij opzet onderzoek: definieer je centrale variabelen en je indicator → essentieel voor methode dataverzameling

# 2. Centrale variabelen en indicator - voorbeelden

Risicofactoren voor slechte *compliance* malariaprofylaxe

- **Centrale variabelen:**
  - Profylaxe ja/nee (recept opgehaald, pillen meegenomen, gestart met profylaxe, kuur volledig afgemaakt)
  - Risicofactoren: leeftijd, geslacht, land bestemming, doel reis (vakantie/werk), duur reis
- **Indicator: RR**

*Vlekjes-ziekte Toolkit* & meldingsgedrag vlekjesziekten KDVs

- **Centrale variabelen**:
  - Toolkit wel/niet,
  - Uitbraken: melding gemaakt uitbraak ja/nee; aantal uitbraken per KDV
- **Indicator**: verschil in percentage KDV dat meldt. Of relatieve maat zoals RR?

Attitude GGD-IZB-medewerkers en CB-medewerkers opname Varicella in RVP

- **Centrale variabelen**:
  - Attitude t.a.v. opname Varicella vaccinatie in RVP
  - Verklarende variabelen: inschatting ernst waterpokken, ….
- **Indicator**: OR

# 3. Definitie en selectie onderzoekspopulatie

- Onderzoekspopulatie = groep waarover je onderzoeksvraag gaat: doelpopulatie
- De onderzoekspopulatie hangt samen met je 'casus-definitie'
  - Risicofactoren infectie hepatitis C drugsgebruikers in Amsterdam in periode X
  - Case = Amsterdamse drugsgebruiker met hepatitis C in periode X
  - Studiepopulatie: alle Amsterdamse drugsgebruikers in periode X
- Selectie onderzoekspopulatie die je includeert hangt af van selectiecriteria, selectiemethode en benodigde sample size

Voorbeelden:

- Risicofactoren voor slechte *compliance* malariaprofylaxe (in periode X)
  - Onderzoekspopulatie: alle reizigers die advies malariaprofylaxe hebben gekregen (in periode X)
- *Vlekjes-ziekte Toolkit* & meldingsgedrag vlekjesziekten KDVs in Arnhem
  - Onderzoekspopulatie: alle KDV in regio Arnhem
- Attitude GGD-IZB-medewerkers en CB-medewerkers opname Varicella in RVP
  - Onderzoekspopulatie: GGD-IZB-medewerkers en CB-medewerkers

# 4. Analyseplan

- Voordat je met studie begint moet je al nadenken over data-analyse
- Wat doe je:
  - Beschrijving van berekeningen en statistische vergelijkingen die je wilt gaan maken
  - Maken van lege tabellen (Dummy tabellen)

- Redenen hiervoor zijn:
  - Focus op de vraagstelling(en) van het onderzoek
  - Zorgt ervoor dat je daadwerkelijk de juiste data verzamelt
  - … en geen onnodige data verzamelt
  - Voorkómen van analyses waarvoor studie niet was opgezet
  - Zorgt ervoor dat je data straks ‘analyseerbaar’ is
    - “Anders, namelijk ....” ?? Levert mogelijk groepen op die je niet kunt analyseren; andere formulering?
  - Tijdsbesparing!!!
    - Het invullen van dummy tabellen gaat snel
    - Het eigenlijke denkwerk zit er al op
    - Minder kans op afleiding en ‘door de bomen het bos niet meer zien’

- Tip: overleg met een expert, epidemioloog, statisticus

# 5. Benodigde data elementen

**Uitganspunt** bij dataverzamelingsplan: welk type informatie heb ik nodig?

- Als het goed is weet je inmiddels hoe je analyse er ongeveer uit gaat zien en dus wat voor data je nodig hebt.
- Je geplande analyse bepaald de data die je moet verzamelen, niet andersom

- Benodigde data elementen:
  1. Demografische gegevens (persoonskenmerken, huishouden)
  2. Centrale variabele: bijvoorbeeld klinische gegevens (bv: ziek ja/nee)
  3. Centrale variabelen: Risicofactoren
  4. Andere factoren: Andere risicofactoren, incl. mogelijke confounders, indicatoren SES

- Data kunnen bestaan uit:
  - Feiten (metingen, labwaarden)
  - Meningen: opinies, attituden
  - Indicatoren van sociaal economische status
  - Lab uitslagen
  - Omgevingsmonsters

# 6. Methode data verzameling

De informatie voor het onderzoek kan afkomstig zijn uit verschillende bronnen:

- Metingen: labtesten, lengte, gewicht, bloeddruk
- Bestaande databronnen: surveillance data, EPD data, andere registraties (vaccinatie RIVM)
- Vragenlijsten
- Andere data:
  - Sociaal economische data via postcode (indicatoren)
- Overig: apotheek recept profylaxe opgehaald

- Alle info per onderzoekspersoon komt op 1 dataverzamelingsformulier (op papier of digitaal) Dit moet uiteindelijk worden samengevoegd in een database
  - NB: ID nummer voor elke persoon, handig voor het samenvoegen van data

**Voorbeeld**

- Malaria profylaxe studie
  - Elementen: demografische factoren, profylaxe gebruik, risicofactoren
  - Allen te verzamelen middels vragenlijst (online of telefonisch?) na terugkomst reis
- KDV en vlekjesziekten:
  - Elementen: ‘demografie’ KDV, meldingen gemaakt van uitbraken, toolkit ja/nee, overige factoren
  - Methoden: ‘interview’ met KDV demografie, overige factoren. Eigen registratie: toolkit ja/nee. Uitbraken gemeld aan GGD

# 7. Uitwerking dataverzameling VRAGENLIJST

Voordelen vragenlijst:

- Je kunt een grote onderzoekspopulatie bereiken
- Relatief makkelijk en effectief/goedkoop
- Kwantificeerbare of meetbare antwoorden op je vragen
- Relatief makkelijk te analyseren (mits goed opgesteld)

Nadelen

- Geeft een eenzijdig inzicht in een probleem
  - Zeker bij beperkte antwoord-opties
  - Vragen kunnen misleidend zijn
- Onduidelijke vragen kunnen zorgen voor: misverstanden / misinterpretatie
- Weinig mogelijkheid tot correctie als vragenlijst eenmaal is uitgezet
- Mogelijke bias (geldt niet alleen voor vragenlijsten).
  - Te reduceren door: gestructureerde vragenlijst & goede respons

# Een goed ontworpen vragenlijst:

- Ziet er mooi uit – prettige lay-out
- Is kort en simpel
- Correcte nummering en verwijzing
- Relevant en logisch: doel van studie is helder voor respondenten
- Begeleidende brief: wie ben je, waar werk je, wat onderzoek je, hoe kom je aan contactgegevens van deze persoon, garandeerd *confidentiality,* tijdsduur (eerlijk!)

- Eenduidige vragen
  - Heb je hamburgers gegeten bij de BBQ?: ☐ veel ☐weinig
  - Hoeveel hamburgers heb je gegeten bij de BBQ? ☐Geen ☐1 ☐2 ☐3 of meer
- Een vraag per vraag
  - Ben je op de BBQ geweest en heb je hamburgers, kip of saus gegeten? *Wat betekent 'ja'*

# Open vragen vs. gesloten vragen

Open vragen: afhankelijk van je studietype en doel

- Voordelen
  - Exploratie fase; hypotheses genereren
  - Nuttig om attitutes en meningen te inventariseren (Qualitatieve studie/ focus groepen)
- Nadelen: Interviewer bias, Kost veel tijd, Lastig te coderen, analyseren

Gesloten vragen:

- Voordelen: Simpel, Makkelijk te coderen, invoeren, analyseren, vergelijken
- Nadelen: Beperkt aantal antwoord opties, Mogelijk verlies van informatie

Optie: mix van gesloten en open vragen

Tot slot - vragenlijst:

- Uitgangspunt = jouw eigen onderzoek en vraagstelling, centrale variabelen
- Uitgangspunt = niet: vragenlijst van iemand anders! Want:
  - Deze is niet opgesteld voor jouw vraagstelling
  - Grote kans op vergeten verzamelen van essentiële informatie
  - Grote kans op verzamelen onnodige informatie
- Vergeet niet je onderzoekspersonen te bedanken en te vertellen waar uitkomsten van onderzoek te vinden zijn

# 8. Dataverzamelingsplan

- Hoe kom  je aan de data die je nodig hebt
  - Vragenlijsten/interview
  - Labs
  - Registraties
- Hoe voeg je data uit verschillende bronnen samen
  - Nodig: ID-nummer
- Trainen van dataverzamelingsteam / interviewers
  - Afspraken: uniforme methode dataverzameling
- Anonieme dataverwerking
  - Labeling, codering
- Online data-entry tools beschikbaar, bv Epi DataEntry (http://www.epidata.dk/)

- Regels en wetgeving:
  - Goedkeuring Medisch Ethische Toetsings Commissie
    - http://www.ccmo-online.nl/main.asp
  - Code Goed Gedrag / Code Goed Gebruik; o.a. verzamelen, bewaren, verwerken persoonsgegeven voor medisch onderzoek
    - http://www.federa.org/gedragscodes-codes-conduct-en

# Samenvatting van de voorbeelden

| **Probleem**<br><br>**Vraagstelling** *(NB: niet SMART)* | Mensen met **import malaria** geen/slecht profylaxe gebruikt<br>Nodig: aanknopingspunten advies gesprek reizigerspoli<br>*Vraag: risicofactoren slechte compliance malariaprofylaxe* |
|---|---|
| **1. Welk onderzoeksdesign** | Cohort studie |
| **2. Centrale variabelen + Indicator** | -Compliance profylaxe<br>-Risicofactoren<br>RR |
| **3. Onderzoekspopulatie En selectie:** | Reizigers op poli GGD Nijmegen met advies malariaprofylaxe<br>Selectie: periode juni 2012 – januari 2013 |
| **4. Analyseplan** | Beschrijving: %compliance, voor verschillende groepen<br>Statistiek: voorspellende factoren – RR |
| **5. Benodigde data elementen** | -Compliance: definitie 'slechte compliance'<br>-Risicofactoren (persoonskenmerken leeftijd, geslacht; reiskenmerken: bestemming, duur reis, doel reis) |
| **6. Methode data verzameling** | Vragenlijst: telefonisch of online |
| **7. Uitwerking dataverzameling** | Specifieke set met goed geformuleerde vragen |
| **8. Dataverzamelingsplan** | o.a. Training GGD-verpleegkundige en medisch student voor tel.interviews; data invoer via EpiDataEntry |

| **Probleem**<br><br>**Vraagstelling** *(NB: niet SMART)* | **Niet altijd melding gemaakt van vlekjesziekte, dus GGD kan niet altijd helpen bij uitbraakbestrijding.**<br>**Verhoogt *Vlekjes-ziekte Toolkit* het %KDV melding uitbraken?** |
|---|---|
| **1. Welk onderzoeksdesign** | Community trial? (als je toolkit als interventie ziet) |
| **2. Centrale variabelen + Indicator** | Melding uitbraken vleksjesziekten<br>Toolkit ja/nee (en type: incl bezoek of per post) |
| **3. Onderzoekspopulatie En selectie:** | Alle KDV in regio Arnhem |
| **4. Analyseplan** | Beschrijving %KDV dat uitbraken heeft gemeld in onderzoeksperiode, opgesplitst voor toolkit ja/nee<br>OR toolkit? ook OR andere factoren |
| **5. Benodigde data elementen** | Elementen: 'demografie' KDV, meldingen gemaakt van uitbraken, toolkit ja/nee, overige factoren |
| **6. Methode data verzameling** | Methoden: 'interview' met KDV demografie, overige factoren.<br>Eigen registratie: toolkit ja/nee. Uitbraken gemeld aan GGD |
| **7. Uitwerking dataverzameling** | Vragenlijst voor interviews<br>Data entry sheet / dataverzamelingsformulier |
| **8. Dataverzamelingsplan** | o.a. Training GGD-verpleegkundige interviews; …… |

| **Probleem**<br><br>**Vraagstelling** *(NB: niet SMART)* | Info nodig om besluitvorming over opname varicella vaccinatie in RVP te ondersteunen<br>**Attitude van professionals t.a.v. opname van Varicella in RVP en risicofactoren?** |
|---|---|
| **1. Welk onderzoeksdesign** | Cross sectionele studie |
| **2. Centrale variabelen + Indicator** | Attitude t.a.v. opname Varicella vaccinatie in RVP<br>Verklarende variabelen: inschatting ernst waterpokken, …. |
| **3. Onderzoekspopulatie En selectie:** | GGD-IZB-medewerkers en CB-medewerkers in Nederland |
| **4. Analyseplan** | Beschrijving % professionals met positieve/negatieve attitude t.a.v. opname Varicella in RVP; opgesplitst voor verschillende factoren<br>Statistisch: OR |
| **5. Benodigde data elementen** | Attitude t.a.v varicella RVP<br>Demografisch, werk kenmerken; overige factoren, ernst waterp. |
| **6. Methode data verzameling** | Vragenlijst online |
| **7. Uitwerking dataverzameling** | Vragen specificeren |
| **8. Dataverzamelingsplan** | |

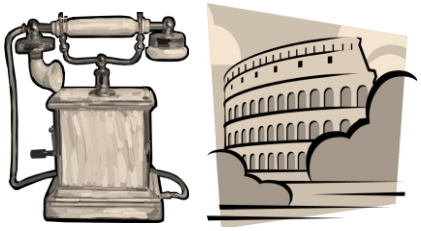
Dit wil je voorkomen!!
1. Start onderzoek
Definitie probleem
2. Opzet onderzoek
Vraagstelling (SMART)
DOEL A
5. Data analyse
Komt volgende keer
3. Data verzameling
Vandaag over gehad
HELP –
Met mijn data kan ik geen antwoord geven op mijn onderzoeksvraag
6. Conclusies & aanbevelingen
4e Tele-ARENA

# Take home messages

- Zorg dat je vraagstelling zo helder mogelijk zijn
  - Revisie van vraagstelling tijdens uitwerking plannen is niet erg maar juist goed → waarborgen kwaliteit van je onderzoek
- Kies de juiste Indicatoren (uitkomstmaat) en Centrale Variabelen
- Bedenk goed hoe je die meetbaar gaat maken
- Maak een analyseplan voordat je met de dataverzameling begint
- Informatie kan uit verschillende databronnen komen
  - Maak een plan voor hoe je de data gaat samenvoegen
- Denk tijdig aan eventuele ethische zaken (METC goedkeuring;  anonieme dataverwerking en veilige opslag)

- Als je deze fase zorgvuldig doorloopt dan zal je
  - Alle data verzamelen die nodig is om onderzoeksvraag te beantwoorden
  - Voorkómen dat je data verzamelt die je niet nodig hebt; zonde van ieders tijd/energie/geld

# Tot slot



- Vragen & suggesties
- Volgende Tele-ARENA: 1 oktober 2013; 12:00u

**Contact**
Alma Tostmann, *infectieziekten epidemioloog*
Universitair Medisch Centrum St Radboud
**Academische Werkplaats AMPHI**
Afdeling Eerstelijnsgeneeskunde
Geert Grooteplein Noord 21, Nijmegen
A.Tostmann@elg.umcn.nl

# Extra info bronnen

- Onderzoeksmethodologie:
  - Boek van *Kenneth Rothman: Modern Epidemiology*
- Sample size berekening –online tools:
  - http://www.stat.ubc.ca/~rollin/stats/ssize/
  - http://www.epibiostat.ucsf.edu/biostat/sampsize.html

Andere online tools voor o.a. sample size maar ook basic epi analyse

- Online EPI: **Open Source Epidemiologic Statistics for Public Health**
  - http://www.openepi.com/v37/Menu/OE_Menu.htm
- Epi-sheet
  - http://krothman.hostbyet2.com/Episheet.xls
- Epidemiology Net
  - http://www.epidemiolog.net/studymat/

**Data entry**: http://www.epidata.dk/