

Analysis of German notification time: 2/3 of notifications arrive until the next day

J. Schumacher, D. Schumacher, M. Diercke, M. Salmon, I. Czogiel, H. Claus, A. Gilsdorf

04.11.2015

About this document

This document calculates and produces the numbers for a poster on ESCAIDE 2015, Stockholm, Sweden. It analyzes the notification delay in the German Surveillance System

Authors: Jakob Schumacher (1)(2)(3), Michaela Diercke (1), Maelle Salmon (1), Irina Czogiel (1)(2), Claus Herrmann (1), Andreas Gilsdorf (1) Affiliations: (1) Robert Koch Institute, Berlin, (2) PAE (Postgraduierte-nausbildung für angewandte Epidemiologie), (3) EPIET (European Programme for Intervention Epidemiology Training)

Preparation

```
#####  
# Setting Path  
#####  
setwd("T:/FG32_Meldeverzug/Projekt_ESCAIDE")  
  
#####  
# Loading packages  
#####  
# Reshaping data  
if(!require("dplyr")) install.packages("dplyr")  
suppressPackageStartupMessages(library("dplyr"))  
# Reshaping data  
if(!require("tidyr")) install.packages("tidyr")  
suppressPackageStartupMessages(library("tidyr"))  
# Graphics  
if(!require("ggplot2")) install.packages("ggplot2")  
suppressPackageStartupMessages(library("ggplot2"))  
# Output  
if(!require("knitr")) install.packages("knitr")  
suppressMessages(library("knitr"))  
# Database  
if(!require("RODBC")) install.packages("RODBC")  
suppressPackageStartupMessages(library("RODBC"))  
# Multivariable model  
if(!require("MASS")) install.packages('MASS')  
suppressPackageStartupMessages(require("MASS"))  
  
#####  
# Setting Options
```

```
#####
# Knitr chunk options
knitr::opts_chunk$set(cache=FALSE, message=FALSE,
                      warning=FALSE, tidy=FALSE, cache=TRUE)
# Dplyr Options to increase the printed lines
options(dplyr.print_max = 1000)
# Select function of MASS and dplyr collide
select <- dplyr::select
```

Getting the dataset

```
#####
# Querying database
#####
# Creating a function for the query
getData <- function () {
  connection <- odbcDriverConnect("driver={SQL Server};server=sesql10;
                                  UID=SQL_SurvNet3_User;PWD=SurvNet3;
                                  Database=SurvNet3RKI2;",
                                  readOnlyOptimize=TRUE)

  query <- "
SELECT

-- Time points --
  CAST(NA1.[DiagnosedAt] AS DATE) AS [ArztDiagnoseAm]
  ,CAST(NA1.[NotifiedAt] AS DATE) AS [ArztMeldungImGA]
  ,CAST(NL1.[DiagnosedAt] AS DATE) AS [LaborDiagnoseAm]
  ,CAST(NL1.[NotifiedAt] AS DATE) AS [LaborMeldungImGA]
  ,CAST(N3.[DiagnosedAt] AS DATE) AS [EigeneErmittlungDiagnoseAm]
  ,CAST(N3.[NotifiedAt] AS DATE) AS [EigeneErmittlungMeldungImGA]
  ,CAST(N4.[DiagnosedAt] AS DATE) AS [GemeinschaftP8DiagnoseAm]
  ,CAST(N4.[NotifiedAt] AS DATE) AS [GemeinschaftP8MeldungImGA]
  ,CAST(N5.[DiagnosedAt] AS DATE) AS [WeiterleitungAnderesGADiagnoseAm]
  ,CAST(N5.[NotifiedAt] AS DATE) AS [WeiterleitungAnderesGAMeldungImGA]
  ,CAST(N6.[DiagnosedAt] AS DATE) AS [GemeinschaftP34DiagnoseAm]
  ,CAST(N6.[NotifiedAt] AS DATE) AS [GemeinschaftP34MeldungImGA]
  ,CAST(N99.[DiagnosedAt] AS DATE) AS [AndereDiagnoseAm]
  ,CAST(N99.[NotifiedAt] AS DATE) AS [AndereMeldungImGA]
  ,CAST(Nx.[DiagnosedAt] AS DATE) AS [UnbekannteMeldeArtDiagnoseAm]
  ,CAST(Nx.[NotifiedAt] AS DATE) AS [UnbekannteMeldeArtImGA]
  ,CAST(DISEASE71.[ReportingDate] AS DATE) AS [Meldedatum]

-- Disease --
  ,M.[GuiText] AS [Disease]

FROM
  [Data].[Version] AS V
  INNER JOIN [Data].[Disease71] AS DISEASE71 ON V.[IdVersion] = DISEASE71.[IdVersion]
  INNER JOIN Meta.DayTable RD ON RD.[IdDaySQL] = CAST(CAST(DISEASE71.[ReportingDate]
AS FLOAT) AS INT)
  INNER JOIN Meta.[Type] M ON V.IdType = M.IdType
```

```

INNER JOIN Meta.[Disease] MD ON V.IdType = MD.IdType AND V.IdSchema = MD.IdSchema
OUTER APPLY [Data].[ExpandWithVaccination] (V.[IdVersion]) Vacc
OUTER APPLY [Data].[ExpandWithDiseasePathogen] (V.[IdVersion]) DisPat

OUTER APPLY (SELECT NF.[IdVersion], NF.[NotificationType] FROM
(SELECT N.[IdVersion], N.[NotificationType], ROW_NUMBER() OVER
(PARTITION BY N.[IdVersion] ORDER BY N.[IdDisease71Notification])
AS [ROWNUMBER] FROM [Data].[Disease71Notification] N WHERE
V.[IdVersion] = N.[IdVersion]) NF WHERE NF.[ROWNUMBER] = 1) NF1

OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification]
N WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 1 GROUP BY [IdVersion]) NA1
OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification]
N WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 2 GROUP BY [IdVersion]) NL1
OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification]
N WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 3 GROUP BY [IdVersion]) N3
OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification] N
WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 4 GROUP BY [IdVersion]) N4
OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification] N
WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 5 GROUP BY [IdVersion]) N5

OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification] N
WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 6 GROUP BY [IdVersion]) N6
OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification] N
WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] = 999999 GROUP BY
[IdVersion]) N99
OUTER APPLY (SELECT [IdVersion], MIN([NotifiedAt]) AS [NotifiedAt],
MIN([DiagnosedAt]) AS [DiagnosedAt] FROM [Data].[Disease71Notification] N
WHERE V.[IdVersion] = N.[IdVersion] AND [NotificationType] IN (0,-1) GROUP BY
[IdVersion]) Nx

OUTER APPLY (SELECT [IdRecord], MIN([TrackedAt]) AS [TrackedAt] FROM
[Data].[RecordTrack] RT WHERE V.[IdRecord] = RT.[IdRecord] AND
[Action] = 1 GROUP BY [IdRecord]) RT1
OUTER APPLY (SELECT [IdRecord], MIN([TrackedAt]) AS [TrackedAt] FROM
[Data].[RecordTrack] RT WHERE V.[IdRecord] = RT.[IdRecord] AND
[Action] = 2 AND [CodeSite] LIKE '1.__._%' GROUP BY [IdRecord]) RT2
OUTER APPLY (SELECT [IdRecord], MIN([TrackedAt]) AS [TrackedAt] FROM
[Data].[RecordTrack] RT WHERE V.[IdRecord] = RT.[IdRecord] AND
[Action] = 3 AND [CodeSite] LIKE '1.__.' GROUP BY [IdRecord]) RT3
OUTER APPLY (SELECT [IdRecord], MIN([TrackedAt]) AS [TrackedAt] FROM
[Data].[RecordTrack] RT WHERE V.[IdRecord] = RT.[IdRecord] AND
[Action] = 2 AND [CodeSite] LIKE '1.__.' GROUP BY [IdRecord]) RT4
OUTER APPLY (SELECT [IdRecord], MIN([TrackedAt]) AS [TrackedAt] FROM
[Data].[RecordTrack] RT WHERE V.[IdRecord] = RT.[IdRecord] AND
[Action] = 3 AND [CodeSite] = '1.' GROUP BY [IdRecord]) RT5

```

```

    OUTER APPLY [Data].[ExpandWithSoftware] (V.[IdRecord]) SoftwareGA
WHERE
    V.[IdRecordType] = 1    -- nur Faelle
    AND V.[IsCurrent] = 1  -- nur aktuelle Version des Falls
    AND V.[IsActive] = 1   -- nur aktive Faelle (nicht geloescht oder verworfen)
    AND DISEASE71.[ReportingDate] BETWEEN '2011-09-29' AND '2014-09-28'
    "
data <- sqlQuery(connection, query)
close(connection)
data
}

# execute function
#data <- getData()

# Change data type
#data <- data %>% mutate_each(funs(as.Date(., format = "%Y-%m-%d")), 1:17)

# Saving the dataset
#save(data, file="rawData_ESCAIDE.RData")

# Loading file out of the directory
load(file="rawData_ESCAIDE.RData")

# Show structure of data
names(data)

```

```

## [1] "ArztDiagnoseAm"
## [2] "ArztMeldungImGA"
## [3] "LaborDiagnoseAm"
## [4] "LaborMeldungImGA"
## [5] "EigeneErmittlungDiagnoseAm"
## [6] "EigeneErmittlungMeldungImGA"
## [7] "GemeinschaftP8DiagnoseAm"
## [8] "GemeinschaftP8MeldungImGA"
## [9] "WeiterleitungAnderesGADiagnoseAm"
## [10] "WeiterleitungAnderesGAMeldungImGA"
## [11] "GemeinschaftP34DiagnoseAm"
## [12] "GemeinschaftP34MeldungImGA"
## [13] "AndereDiagnoseAm"
## [14] "AndereMeldungImGA"
## [15] "UnbekannteMeldeArtDiagnoseAm"
## [16] "UnbekannteMeldeArtImGA"
## [17] "Meldedatum"
## [18] "Disease"

```

Preparation of the dataset

```

#####
# Restricting diseases to §6.1.1 and §7.1 infection protection act
#####

```

```

DiseaseData <- data.frame(rbind(
##### Paragraph 6.1.1. #####
# a) Botulismus
c(Disease="Botulismus", Meldepflicht="67"),
# b) Cholera
c(Disease="Cholera", Meldepflicht="67"),
# c) Diphtherie
c(Disease="Diphtherie", Meldepflicht="67"),
# d) humaner spongiformer Enzephalopathie, außer familiär-hereditärer Formen
c(Disease="CJK", Meldepflicht="6"),
# e) akuter Virushepatitis
c(Disease="Hepatitis A", Meldepflicht="67"),
c(Disease="Hepatitis B", Meldepflicht="67"),
c(Disease="Hepatitis C", Meldepflicht="67"),
c(Disease="Hepatitis D", Meldepflicht="67"),
c(Disease="Hepatitis E", Meldepflicht="67"),
# f) enteropathischem hämolytisch-urämischem Syndrom (HUS)
c(Disease="HUS", Meldepflicht="6"),
# g) virusbedingtem hämorrhagischen Fieber
c(Disease="Ebolavirus", Meldepflicht="67"),
c(Disease="Lassavirus", Meldepflicht="67"),
c(Disease="Marburgvirus", Meldepflicht="67"),
c(Disease="Denguefieber", Meldepflicht="67"),
c(Disease="VHF, andere Erreger", Meldepflicht="67"),
# h) Masern
c(Disease="Masern", Meldepflicht="67"),
# i) Meningokokken-Meningitis oder -Sepsis
c(Disease="Meningokokken", Meldepflicht="67"),
# j) Milzbrand
c(Disease="Milzbrand", Meldepflicht="67"),
# k) Mumps
c(Disease="Mumps", Meldepflicht="67"),
# l) Pertussis
c(Disease="Keuchhusten", Meldepflicht="67"),
# m) Poliomyelitis (als Verdacht gilt jede akute schlaffe Lähmung)
c(Disease="Poliomyelitis", Meldepflicht="67"),
# n) Pest
c(Disease="Pest", Meldepflicht="67"),
# o) Röteln einschließlich Rötelnembryopathie
c(Disease="Röteln", Meldepflicht="67"),
c(Disease="Röteln, konnatal", Meldepflicht="67"),
# p) Tollwut
c(Disease="Tollwut", Meldepflicht="67"),
# q) Typhus abdominalis/Paratyphus
c(Disease="Typhus", Meldepflicht="67"),
c(Disease="Paratphyus", Meldepflicht="67"),
# r) Varizellen
c(Disease="Windpocken", Meldepflicht="67"),
# sowie die Erkrankung und der Tod an einer behandlungsbedürftigen Tuberkulose
c(Disease="Tuberkulose", Meldepflicht="67"),
##### Paragraph 7.1. #####
# 1. Adenoviren; Meldepflicht nur für den direkten Nachweis im Konjunktivalabstrich
c(Disease="Adenovirus", Meldepflicht="7"),

```

```

# 2. Bacillus anthracis - Already in §6.1.1
# 3. Bordetella pertussis, Bordetella parapertussis - Already in §6.1.1
# 4. Borrelia recurrentis
c(Disease="Läuserückfallfieber", Meldepflicht="7"),
# 5. Brucella sp.
c(Disease="Brucellose", Meldepflicht="7"),
# 6. Campylobacter sp., darmpathogen
c(Disease="Campylobacter", Meldepflicht="7"),
# 7. Chlamydia psittaci
c(Disease="Ornithose", Meldepflicht="7"),
# 8. Clostridium botulinum oder Toxinnachweis - Already in §6.1.1
# 9. Corynebacterium diphtheriae, Toxin bildend - Already in §6.1.1
# 10. Coxiella burnetii
c(Disease="Q-Fieber", Meldepflicht="7"),
# 11. humanpathogene Cryptosporidium sp.
c(Disease="Kryptosporidiose", Meldepflicht="7"),
# 12. Ebolavirus - Already in §6.1.1
# 13. a) Escherichia coli, enterohämorrhagische Stämme (EHEC)
c(Disease="EHEC/STEC", Meldepflicht="7"),
# b) Escherichia coli, sonstige darmpathogene Stämme
c(Disease="E.-coli-Enteritis", Meldepflicht="7"),
# 14. Francisella tularensis
c(Disease="Tularämie", Meldepflicht="7"),
# 15. FSME-Virus
c(Disease="FSME", Meldepflicht="7"),
# 16. Gelbfiebertivirus
c(Disease="Gelbfieber", Meldepflicht="7"),
# 17. Giardia lamblia
c(Disease="Giardiasis", Meldepflicht="7"),
# 18. Haemophilus influenzae; Meldepflicht nur für den direkten Nachweis
c(Disease="Haemophilus influenzae", Meldepflicht="7"),
# 19. Hantaviren
c(Disease="Hantavirus", Meldepflicht="7"),
# 20. Hepatitis-A-Virus - Already in §6.1.1
# 21. Hepatitis-B-Virus - Already in §6.1.1
# 22. Hepatitis-C-Virus; - Already in §6.1.1
# 23. Hepatitis-D-Virus - Already in §6.1.1
# 24. Hepatitis-E-Virus - Already in §6.1.1
# 25. Influenzaviren; Meldepflicht nur für den direkten Nachweis
c(Disease="Influenza", Meldepflicht="7"),
# 26. Lassavirus - Already in §6.1.1
# 27. Legionella sp.
c(Disease="Legionellose", Meldepflicht="7"),
# 28. humanpathogene Leptospira sp.
c(Disease="Leptospirose", Meldepflicht="7"),
# 29. Listeria monocytogenes; Meldepflicht nur für den direkten Nachweis
c(Disease="Listeriose", Meldepflicht="7"),
# 30. Marburgvirus - Already in §6.1.1
# 31. Masernvirus - Already in §6.1.1
# 32. Mumpsvirus - Already in §6.1.1
# 33. Mycobacterium leprae
c(Disease="Lepra", Meldepflicht="7"),
# 34. Mycobacterium tuberculosis/africanum, Mycobacterium bovis; - Already in §6.1.1

```

```

# 35. Neisseria meningitidis; - Already in §6.1.1
# 36. Norwalk-ähnliches Virus;
c(Disease="Norovirus", Meldepflicht="7"),
# 37. Poliovirus - Already in §6.1.1
# 38. Rabiesvirus - Already in §6.1.1
# 39. Rickettsia prowazekii
c(Disease="Fleckfieber", Meldepflicht="7"),
# 40. Rotavirus
c(Disease="Rotavirus", Meldepflicht="7"),
# 41. Rubellavirus - Already in §6.1.1
# 42. Salmonella Paratyphi - Already in §6.1.1
# 43. Salmonella Typhi - Already in §6.1.1
# 44. Salmonella, sonstige
c(Disease="Salmonellose", Meldepflicht="7"),
# 45. Shigella sp.
c(Disease="Shigellose", Meldepflicht="7"),
# 46. Trichinella spiralis
c(Disease="Trichinellose", Meldepflicht="7"),
# 47. Varizella-Zoster-Virus - Already in §6.1.1
# 48. Vibrio cholerae O 1 und O 139 - Already in §6.1.1
# 49. Yersinia enterocolitica, darmpathogen
c(Disease="Yersiniose", Meldepflicht="7"),
# 50. Yersinia pestis
c(Disease="Pest", Meldepflicht="7")
# 51. andere Erreger hämorrhagischer Fieber. - Already in §6.1.1
# Diseases not included:
# Länderverordnung: Adenovirus, Länderverordnung, Amoebiasis,
# Gasbrand, Kopfläuse, Lyme-Borreliose, Meningoenzephalitis viral,
# Pneumokokken, Scharlach, Weitere Bedrohliche,
# Weitere Bedrohliche (gastro)
# Andere Stelle IfSG: Clostridium difficile, MRSA,
# Tollwutexpositionsverdacht
))

# Show number of cases
dim(data)

```

```
## [1] 1226646      18
```

```

# Actual Filtering
data <- data %>% filter(Disease %in% as.character(DiseaseData$Disease))

# Show number of cases
dim(data)

```

```
## [1] 1070284      18
```

```

# Erase Mumps, Röteln, Windpocken, Keuchhusten because
# they exist only after amendment of the infection protection act
data <- data %>%
  filter(Disease!="Mumps" &
         Disease!="Keuchhusten" &

```

```

Disease!="Röteln" &
Disease!="Windpocken") %>%
droplevels()

# Show number of cases
dim(data)

```

```
## [1] 998613      18
```

```

# Show available diseases
table(data$Disease)

```

```
##
##          Adenovirus          Botulismus          Brucellose
##          6104              14              116
##          Campylobacter          Cholera          CJK
##          208471              1              375
##          Denguefieber          Diphtherie          E.-coli-Enteritis
##          2122              23              25500
##          EHEC/STEC          FSME          Giardiasis
##          6130              943              17127
## Haemophilus influenzae          Hantavirus          Hepatitis A
##          1257              3702              2849
##          Hepatitis B          Hepatitis C          Hepatitis D
##          7439              16968              173
##          Hepatitis E          HUS          Influenza
##          1663              236              94893
##          Kryptosporidiose          Legionellose          Lepra
##          4774              2644              8
##          Leptospirose          Listeriose          Masern
##          326              1627              2395
##          Meningokokken          Milzbrand          Norovirus
##          1013              4              370281
##          Ornithose          Poliomyelitis          Q-Fieber
##          45              2              822
##          Rotavirus          Röteln, konnatal          Salmonellose
##          131277              2              63584
##          Shigellose          Trichinellose          Tuberkulose
##          1749              33              13010
##          Tularämie          Typhus          VHF, andere Erreger
##          58              210              140
##          Yersiniose
##          8503

```

```

#####
# Calculate additional variables
#####
# ImportGA.
data <- data %>%
  mutate(ImportGA = pmin(ArztMeldungImGA,
                        LaborMeldungImGA,
                        EigeneErmittlungMeldungImGA,

```



```

        GemeinschaftP8MeldungImGA,
        WeiterleitungAnderesGAMeldungImGA,
        GemeinschaftP34MeldungImGA,
        AndereMeldungImGA,
        UnbekannteMeldeArtImGA, na.rm=TRUE)) %>%
mutate(ImportGA = as.Date(ImportGA, origin="1970-01-01"))

# Diagnosis
data <- data %>%
  mutate(Diagnosis = pmin(ArztDiagnoseAm,
    LaborDiagnoseAm,
    EigeneErmittlungDiagnoseAm,
    GemeinschaftP8DiagnoseAm,
    WeiterleitungAnderesGADiagnoseAm,
    GemeinschaftP34DiagnoseAm,
    AndereDiagnoseAm,
    UnbekannteMeldeArtDiagnoseAm, na.rm=TRUE)) %>%
  mutate(Diagnosis = as.Date(Diagnosis, origin="1970-01-01"))

# delays
data <- data %>%
  mutate(delay_arzt = as.numeric(ArztMeldungImGA-ArztDiagnoseAm)) %>%
  mutate(delay_labor = as.numeric(LaborMeldungImGA-LaborDiagnoseAm)) %>%
  mutate(delay_notification = as.numeric(ImportGA - Diagnosis))

# gesetz
data <- data %>%
  mutate(gesetz_diagnosis = factor(ifelse(!is.na(Diagnosis),
    ifelse(Diagnosis < as.Date("2013-03-29"), "VorGesetz", "NachGesetz"),
    ifelse(Melddatum < as.Date("2013-03-29"), "VorGesetz", "NachGesetz"))))

# paragraph
DiseaseData$Disease <- factor(DiseaseData$Disease)
data <- data %>% left_join(DiseaseData, by="Disease") %>% mutate(Disease=factor(Disease))

# notifier (set at first arrival at GA)
data <- data %>%
  mutate(Melder="Other") %>%
  mutate(Melder = ifelse(!is.na(LaborMeldungImGA) &
    LaborMeldungImGA == ImportGA , "Labor", Melder)) %>%
  mutate(Melder = ifelse(!is.na(ArztMeldungImGA) &
    ArztMeldungImGA == ImportGA , "Arzt", Melder)) %>%
  mutate(Melder = ifelse(!is.na(LaborMeldungImGA) &
    LaborMeldungImGA == ImportGA &
    !is.na(ArztMeldungImGA) &
    ArztMeldungImGA == ImportGA, "Beide", Melder)) %>%
  mutate(Melder = ifelse(is.na(ImportGA), "None", Melder)) %>%
  mutate(Melder = factor(Melder))

# Show structure of data
names(data)

```

```
## [1] "ArztDiagnoseAm"
```

```
## [2] "ArztMeldungImGA"
## [3] "LaborDiagnoseAm"
## [4] "LaborMeldungImGA"
## [5] "EigeneErmittlungDiagnoseAm"
## [6] "EigeneErmittlungMeldungImGA"
## [7] "GemeinschaftP8DiagnoseAm"
## [8] "GemeinschaftP8MeldungImGA"
## [9] "WeiterleitungAnderesGADiagnoseAm"
## [10] "WeiterleitungAnderesGAMeldungImGA"
## [11] "GemeinschaftP34DiagnoseAm"
## [12] "GemeinschaftP34MeldungImGA"
## [13] "AndereDiagnoseAm"
## [14] "AndereMeldungImGA"
## [15] "UnbekannteMeldeArtDiagnoseAm"
## [16] "UnbekannteMeldeArtImGA"
## [17] "Meldedatum"
## [18] "Disease"
## [19] "ImportGA"
## [20] "Diagnosis"
## [21] "delay_arzt"
## [22] "delay_labor"
## [23] "delay_notification"
## [24] "gesetz_diagnosis"
## [25] "Meldepflicht"
## [26] "Melder"
```

```
#####
# Restricting to time
#####
# Show number of cases
dim(data)
```

```
## [1] 998613    26
```

```
# Filtering
data <- data %>%
  filter(is.na(Diagnosis) |
         Diagnosis >= as.Date("2012-03-29") &
         Diagnosis <= as.Date("2014-03-28"))

# Show number of cases
dim(data)
```

```
## [1] 758663    26
```

```
#####
# Deleting delay values that are potentially wrong
#####

## Capturing deleted values
sum(!is.na(data %>% filter(grepl(6, Meldepflicht))))
```

```
## [1] 364256
```

```
sum(!is.na(data %>% filter(grepl(6, Meldepflicht)) %>% select(delay_arzt))) /  
  nrow(data %>% filter(grepl(6, Meldepflicht)))
```

```
## [1] 0.1777729
```

```
sum(!is.na(data %>% filter(grepl(7, Meldepflicht))))
```

```
## [1] 7485575
```

```
sum(!is.na(data %>% filter(grepl(7, Meldepflicht)) %>% select(delay_labor))) /  
  nrow(data %>% filter(grepl(7, Meldepflicht)))
```

```
## [1] 0.5779354
```

```
### Far-off values
```

```
data <- data %>%
```

```
  mutate(delay_arzt = ifelse(delay_arzt<0|delay_arzt>183, NA, delay_arzt)) %>%
```

```
  mutate(delay_labor = ifelse(delay_labor<0|delay_labor>183, NA, delay_labor)) %>%
```

```
  mutate(delay_notification = ifelse(delay_notification<0|  
    delay_notification>183, NA, delay_notification))
```

```
## Capturing deleted values
```

```
sum(!is.na(data %>% filter(grepl(6, Meldepflicht))))
```

```
## [1] 362964
```

```
sum(!is.na(data %>% filter(grepl(6, Meldepflicht)) %>% select(delay_arzt))) /  
  nrow(data %>% filter(grepl(6, Meldepflicht)))
```

```
## [1] 0.1719438
```

```
sum(!is.na(data %>% filter(grepl(7, Meldepflicht))))
```

```
## [1] 7476342
```

```
sum(!is.na(data %>% filter(grepl(7, Meldepflicht)) %>% select(delay_labor))) /  
  nrow(data %>% filter(grepl(7, Meldepflicht)))
```

```
## [1] 0.5762049
```

```
# Erase second arrival at GA entries
```

```
data <- data %>%
```

```
  mutate(delay_arzt = ifelse(Melder=="Labor", NA, delay_arzt)) %>%
```

```
  mutate(delay_labor = ifelse(Melder=="Arzt", NA, delay_labor))
```

```
## Capturing deleted values
```

```
sum(!is.na(data %>% filter(grepl(6, Meldepflicht))))
```

```
## [1] 361581
```

```
sum(!is.na(data %>% filter(grepl(6, Meldepflicht)) %>% select(delay_arzt))) /  
  nrow(data %>% filter(grepl(6, Meldepflicht)))
```

```
## [1] 0.1625845
```

```
sum(!is.na(data %>% filter(grepl(7, Meldepflicht))))
```

```
## [1] 7469802
```

```
sum(!is.na(data %>% filter(grepl(7, Meldepflicht)) %>% select(delay_labor))) /  
  nrow(data %>% filter(grepl(7, Meldepflicht)))
```

```
## [1] 0.5713538
```

Analyzing the dataset

```
#####  
# Numbers of notifications physicians  
#####  
  
# Create subset of dataset  
p6 <- data %>% filter(grepl(6, Meldepflicht))  
  
#Numbers  
# Total number of notifications in p6 diseases  
P6Notifications <- p6 %>% nrow  
  
# Number of physicians notifications in p6 diseases  
PhysiciansP6Notification <- p6 %>% filter(!is.na(delay_arzt)) %>% nrow  
  
# Number of physicians notifications in p6 diseases  
LabP6Notification <- p6 %>% filter(!is.na(delay_labor)) %>% nrow  
  
# Number of unkown delay in p6 diseases  
UnkownDelayP6Notification <- p6 %>% filter(is.na(ImportGA)|is.na(Diagnosis)) %>% nrow  
  
# Percentages  
# Percentage of physicians notifications in p6 diseases  
PercentagePhysiciansP6Notification <-  
  round(100*PhysiciansP6Notification/P6Notifications,0)  
  
# Percentage of physicians notifications in p6 diseases  
PercentageLabP6Notification <-  
  round(100*LabP6Notification/P6Notifications,0)  
  
# Percentage of physicians notifications in p6 diseases  
PercentageUnkownDelayP6Notification <-  
  round(100*UnkownDelayP6Notification/P6Notifications,0)
```

```
#####
# Numbers of notifications laboratories
#####

# Create subset of dataset
p7 <- data %>% filter(grepl(7, Meldepflicht))

#Numbers
# Total number of notifications in p7 diseases
P7Notifications <- p7 %>% nrow

# Number of physicians notifications in p7 diseases
PhysiciansP7Notification <- p7 %>% filter(!is.na(delay_arzt)) %>% nrow

# Number of physicians notifications in p7 diseases
LabP7Notification <- p7 %>% filter(!is.na(delay_labor)) %>% nrow

# Number of unkown delay in p7 diseases
UnkownDelayP7Notification <- p7 %>%
  filter(is.na(ImportGA)|is.na(Diagnosis)) %>% nrow

# Percentages
# Percentage of physicians notifications in p7 diseases
PercentagePhysiciansP7Notification <-
  round(100*PhysiciansP7Notification/P7Notifications,0)

# Percentage of physicians notifications in p7 diseases
PercentageLabP7Notification <-
  round(100*LabP7Notification/P7Notifications,0)

# Percentage of physicians notifications in p7 diseases
PercentageUnkownDelayP7Notification <-
  round(100*UnkownDelayP7Notification/P7Notifications,0)

# Display numbers
rbind(P6Notifications,
      PhysiciansP6Notification,
      LabP6Notification,
      UnkownDelayP6Notification,
      PercentagePhysiciansP6Notification,
      PercentageLabP6Notification,
      PercentageUnkownDelayP6Notification,
      P7Notifications,
      PhysiciansP7Notification,
      LabP7Notification,
      UnkownDelayP7Notification,
      PercentagePhysiciansP7Notification,
      PercentageLabP7Notification,
      PercentageUnkownDelayP7Notification)

##                               [,1]
## P6Notifications              36541
## PhysiciansP6Notification     5941
```

```
## LabP6Notification          18894
## UnkownDelayP6Notification  11664
## PercentagePhysiciansP6Notification    16
## PercentageLabP6Notification    52
## PercentageUnkownDelayP6Notification    32
## P7Notifications          758173
## PhysiciansP7Notification    65546
## LabP7Notification          433185
## UnkownDelayP7Notification  254224
## PercentagePhysiciansP7Notification    9
## PercentageLabP7Notification    57
## PercentageUnkownDelayP7Notification    34
```

Tables

```
#####
# Calculations
#####
# Percentage within days overall
kable(
  data %>%
  filter(!is.na(delay_notification)) %>%
  summarise(GesamtAnzahl = n(),
            InnerhalbNullTageA = round(sum(delay_notification==0)*100/GesamtAnzahl,1),
            InnerhalbEinsTageA = round(sum(delay_notification<=1)*100/GesamtAnzahl,1),
            InnerhalbZweiTageA = round(sum(delay_notification<=2)*100/GesamtAnzahl,1),
            InnerhalbDreiTageA = round(sum(delay_notification<=3)*100/GesamtAnzahl,1),
            InnerhalbVierTageA = round(sum(delay_notification<=4)*100/GesamtAnzahl,1),
            InnerhalbFuenfTageA = round(sum(delay_notification<=5)*100/GesamtAnzahl,1)) %>%
  select(-GesamtAnzahl) %>%
  gather()
)
```

key	value
InnerhalbNullTageA	45.3
InnerhalbEinsTageA	66.1
InnerhalbZweiTageA	75.8
InnerhalbDreiTageA	85.5
InnerhalbVierTageA	90.9
InnerhalbFuenfTageA	94.0

```
# Alternative way:
# P <- ecdf(data$delay_notification)
# P(1:10)

# Percentage within days overall grouped before and after the amendment
kable(data %>%
  filter(!is.na(delay_notification)) %>%
  group_by(gesetz_diagnosis) %>%
```

```

summarise(GesamtAnzahl = n(),
  InnerhalbNullTageA = round(sum(delay_notification==0)*100/GesamtAnzahl,1),
  InnerhalbEinsTageA = round(sum(delay_notification<=1)*100/GesamtAnzahl,1),
  InnerhalbZweiTageA = round(sum(delay_notification<=2)*100/GesamtAnzahl,1),
  InnerhalbDreiTageA = round(sum(delay_notification<=3)*100/GesamtAnzahl,1),
  InnerhalbVierTageA = round(sum(delay_notification<=4)*100/GesamtAnzahl,1),
  InnerhalbFuenfTageA = round(sum(delay_notification<=5)*100/GesamtAnzahl,1)) %>%
select(-GesamtAnzahl) %>%
  gather(var,val, 2:7) %>%
  spread(gesetz_diagnosis, val) %>%
  select(var, VorGesetz, NachGesetz)
)

```

var	VorGesetz	NachGesetz
InnerhalbNullTageA	43.6	47.4
InnerhalbEinsTageA	65.1	67.5
InnerhalbZweiTageA	75.0	76.8
InnerhalbDreiTageA	84.6	86.7
InnerhalbVierTageA	90.2	91.9
InnerhalbFuenfTageA	93.5	94.7

```

#####
# Disease stratification
#####
# Measles
P <- data %>% filter(Disease=="Masern") %>% .[["delay_notification"]] %>% ecdf(); P(1)

```

```
## [1] 0.6916472
```

```

# Meningokokken
P <- data %>% filter(Disease=="Meningokokken") %>% .[["delay_notification"]] %>% ecdf(); P(1)

```

```
## [1] 0.7896825
```

```

# CJD
P <- data %>% filter(Disease=="CJK") %>% .[["delay_notification"]] %>% ecdf(); P(1)

```

```
## [1] 0.4135338
```

```

# Tuberkulose
P <- data %>% filter(Disease=="Tuberkulose") %>% .[["delay_notification"]] %>% ecdf(); P(1)

```

```
## [1] 0.5419223
```

Graphs

```

#####
# Time series plot
#####
#####
# Time series plot overall
#####
# Set Variables
fromDate <- as.Date("2012-03-29")
toDate <- as.Date("2014-03-28")

# Calculate the data
data$yearweek <- as.numeric(strftime(data$Diagnosis, format="%Y%U"))
graph_data_overall <- data %>%
  filter(Diagnosis>fromDate&Diagnosis<toDate) %>%
  group_by(yearweek) %>%
  summarise(count=n(), mean_delay_overall = mean(delay_notification, na.rm=TRUE))

# Axislabels
# Turn the yearweek variable into a factor
graph_data_overall$yearweek<-factor(graph_data_overall$yearweek)
# Extract the axislabels
axislabels <- paste0(
  substr(as.character(graph_data_overall$yearweek),5,6),
  "-",
  substr(as.character(graph_data_overall$yearweek),1,4))
# Get the interval to equalize axislabels and breaks
interval <- seq(match("00-2013", axislabels),
  length(graph_data_overall$yearweek),53)

# Draw vertical lines
amendment <- match("13-2013", axislabels)

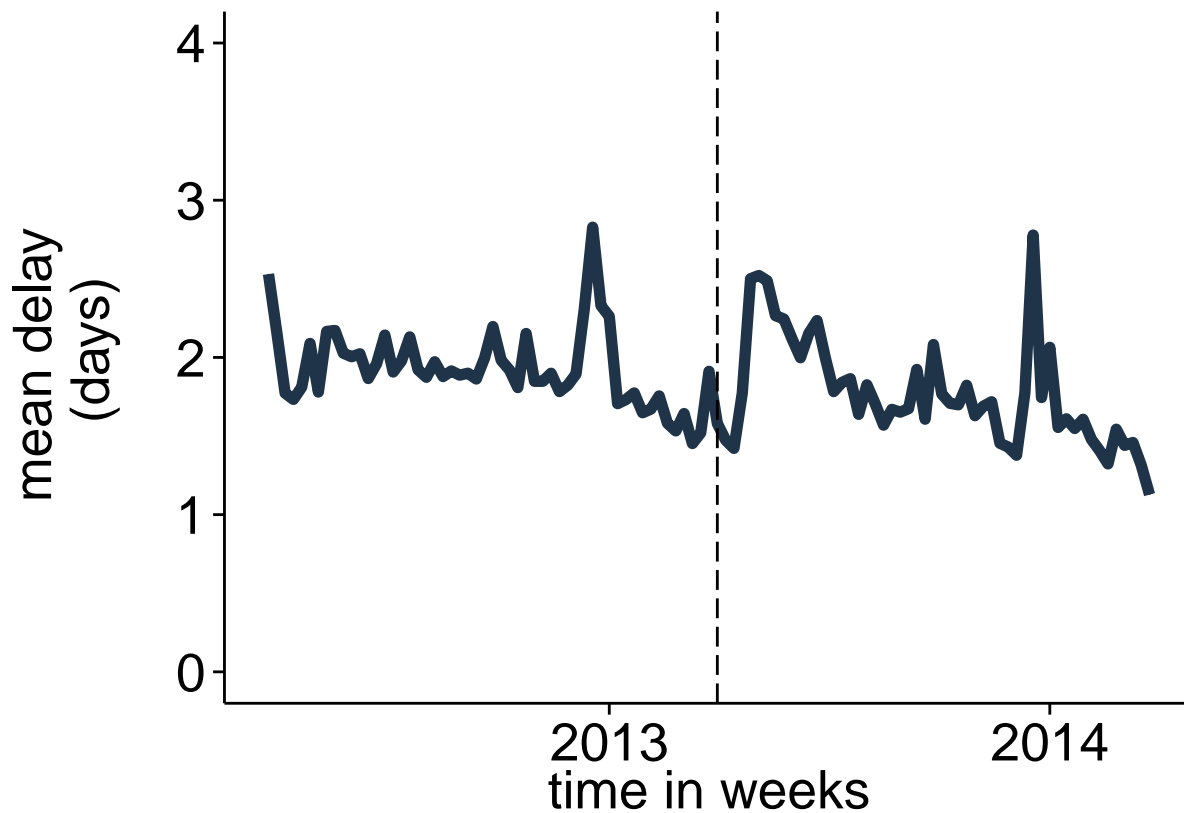
# Turn axislabels into years only
axislabels <- substr(axislabels, 4,7)

# The actual plotting
pOverall <- ggplot(graph_data_overall, aes(as.numeric(yearweek),
  mean_delay_overall)) +
  geom_line() +
  theme_classic() +
  geom_line(color="#1F3449", size=2) +
  geom_vline(aes(xintercept=amendment), linetype = "longdash") +
  scale_x_continuous(name="time in weeks",
    breaks=interval,
    labels=axislabels[interval]) +
  scale_y_continuous(name="mean delay \n (days)\n", limits=c(0,4)) +
  theme(axis.text = element_text(size=20),
    axis.title = element_text(size=20))

# Viewing and saving
ggsave("Timeseries.png", pOverall)

```


pOverall



Multivariable model

```
model<-glm.nb(delay_notification~gesetz_diagnosis + Melder + Disease , data=data)
summary(model)
```

```
##
## Call:
## glm.nb(formula = delay_notification ~ gesetz_diagnosis + Melder +
##   Disease, data = data, init.theta = 0.5500565975, link = log)
##
## Deviance Residuals:
##   Min       1Q   Median       3Q      Max
## -2.0576  -1.1798  -0.4158   0.2423  11.0588
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.923116   0.029057  31.769 < 2e-16 ***
## gesetz_diagnosisVorGesetz 0.077835   0.004606  16.898 < 2e-16 ***
## MelderBeide    0.034838   0.013910   2.505 0.012262 *
## MelderLabor   -0.522066   0.012763 -40.905 < 2e-16 ***
```

```

## MelderOther          -0.210595  0.014244 -14.785 < 2e-16 ***
## DiseaseBotulismus    2.228245  0.612459  3.638 0.000275 ***
## DiseaseBrucellose    0.489985  0.194353  2.521 0.011699 *
## DiseaseCampylobacter -0.015347  0.027136 -0.566 0.571703
## DiseaseCholera       -14.781470 773.785005 -0.019 0.984759
## DiseaseCJK           1.529927  0.123038 12.435 < 2e-16 ***
## DiseaseDenguefieber  0.250262  0.049087  5.098 3.43e-07 ***
## DiseaseDiphtherie    -1.626125  0.604050 -2.692 0.007102 **
## DiseaseE.-coli-Enteritis 0.199985  0.029869  6.695 2.15e-11 ***
## DiseaseEHEC/STEC     0.097540  0.038113  2.559 0.010490 *
## DiseaseFSME          0.036029  0.071493  0.504 0.614295
## DiseaseGiardiasis    0.545547  0.030978 17.611 < 2e-16 ***
## DiseaseHaemophilus influenzae 0.294930  0.062313  4.733 2.21e-06 ***
## DiseaseHantavirus    -0.027815  0.041478 -0.671 0.502478
## DiseaseHepatitis A   -0.023447  0.046878 -0.500 0.616959
## DiseaseHepatitis B    1.110923  0.035112 31.639 < 2e-16 ***
## DiseaseHepatitis C    0.839292  0.030693 27.345 < 2e-16 ***
## DiseaseHepatitis D    0.754599  0.147336  5.122 3.03e-07 ***
## DiseaseHepatitis E    0.528117  0.055909  9.446 < 2e-16 ***
## DiseaseHUS           -0.046415  0.151973 -0.305 0.760051
## DiseaseInfluenza     -0.212064  0.027460 -7.723 1.14e-14 ***
## DiseaseKryptosporidiose -0.023706  0.041031 -0.578 0.563422
## DiseaseLegionellose  0.116725  0.049102  2.377 0.017446 *
## DiseaseLepra         2.480707  0.964550  2.572 0.010115 *
## DiseaseLeptospirose  0.999858  0.115544  8.653 < 2e-16 ***
## DiseaseListeriose    0.240064  0.059035  4.066 4.77e-05 ***
## DiseaseMasern        -0.263852  0.051390 -5.134 2.83e-07 ***
## DiseaseMeningokokken -0.633265  0.079008 -8.015 1.10e-15 ***
## DiseaseMilzbrand     -2.237914  1.243837 -1.799 0.071987 .
## DiseaseNorovirus     -0.165159  0.027014 -6.114 9.72e-10 ***
## DiseaseOrnithose     1.059181  0.310661  3.409 0.000651 ***
## DiseasePoliomyelitis  1.223289  0.993359  1.231 0.218148
## DiseaseQ-Fieber      0.353856  0.084672  4.179 2.93e-05 ***
## DiseaseRotavirus     -0.142800  0.027404 -5.211 1.88e-07 ***
## DiseaseRöteln, konnatal -14.703635 773.785006 -0.019 0.984839
## DiseaseSalmonellose  0.099847  0.027996  3.566 0.000362 ***
## DiseaseShigellose    0.233662  0.057287  4.079 4.53e-05 ***
## DiseaseTrichinellose  1.831803  0.399140  4.589 4.45e-06 ***
## DiseaseTuberkulose   1.191669  0.033538 35.532 < 2e-16 ***
## DiseaseTularämie     0.137309  0.262956  0.522 0.601548
## DiseaseTyphus        0.159330  0.139799  1.140 0.254406
## DiseaseVHF, andere Erreger 0.855510  0.301368  2.839 0.004529 **
## DiseaseYersiniose    0.074590  0.035575  2.097 0.036020 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.5501) family taken to be 1)
##
## Null deviance: 511739 on 496949 degrees of freedom
## Residual deviance: 483150 on 496903 degrees of freedom
## (261713 observations deleted due to missingness)
## AIC: 1730629
##
## Number of Fisher Scoring iterations: 1

```

```
##
##
##           Theta: 0.55006
##          Std. Err.: 0.00176
##
## 2 x log-likelihood: -1730532.88600
```

```
round(exp(model$coefficients),15)
```

```
##           (Intercept)      gesetz_diagnosisVorGesetz
##           2.517122e+00                1.080944e+00
##           MelderBeide                MelderLabor
##           1.035452e+00                5.932936e-01
##           MelderOther                DiseaseBotulismus
##           8.101018e-01                9.283563e+00
##           DiseaseBrucellose          DiseaseCampylobacter
##           1.632292e+00                9.847704e-01
##           DiseaseCholera                DiseaseCJK
##           3.806177e-07                4.617840e+00
##           DiseaseDenguefieber          DiseaseDiphtherie
##           1.284362e+00                1.966903e-01
##           DiseaseE.-coli-Enteritis    DiseaseEHEC/STEC
##           1.221385e+00                1.102455e+00
##           DiseaseFSME                DiseaseGiardiasis
##           1.036686e+00                1.725552e+00
##           DiseaseHaemophilus influenzae DiseaseHantavirus
##           1.343032e+00                9.725681e-01
##           DiseaseHepatitis A          DiseaseHepatitis B
##           9.768259e-01                3.037159e+00
##           DiseaseHepatitis C          DiseaseHepatitis D
##           2.314728e+00                2.126759e+00
##           DiseaseHepatitis E          DiseaseHUS
##           1.695735e+00                9.546461e-01
##           DiseaseInfluenza            DiseaseKryptosporidiose
##           8.089133e-01                9.765725e-01
##           DiseaseLegionellose          DiseaseLepra
##           1.123810e+00                1.194971e+01
##           DiseaseLeptospirose          DiseaseListeriose
##           2.717896e+00                1.271330e+00
##           DiseaseMasern                DiseaseMeningokokken
##           7.680870e-01                5.308559e-01
##           DiseaseMilzbrand            DiseaseNorovirus
##           1.066808e-01                8.477593e-01
##           DiseaseOrnithose            DiseasePoliomyelitis
##           2.884007e+00                3.398345e+00
##           DiseaseQ-Fieber              DiseaseRotavirus
##           1.424550e+00                8.669271e-01
##           DiseaseRöteln, konnatal     DiseaseSalmonellose
##           4.114265e-07                1.105001e+00
##           DiseaseShigellose            DiseaseTrichinellose
##           1.263218e+00                6.245138e+00
##           DiseaseTuberkulose            DiseaseTularämie
##           3.292572e+00                1.147183e+00
##           DiseaseTyphus                DiseaseVHF, andere Erreger
```

##	1.172725e+00	2.352575e+00
##	DiseaseYersiniose	
##	1.077442e+00	