IMPROVING EPIDEMIOLOGICAL STUDIES ON HEALTH EFFECTS FROM RADIOFREQUENCY ELECTROMAGNETIC FIELDS EXPOSURE

Marije Reedijk

Improving epidemiological studies on health effects from radiofrequency electromagnetic fields exposure

Marije Reedijk

2022

ISBN: 978-94-6458-123-2
Cover design: Publiss | www.publiss.nl – photo shutterstock
Lay-out: Publiss | www.publiss.nl
Print: Ridderprint | www.ridderprint.nl

Improving epidemiological studies on health effects from radiofrequency electromagnetic fields exposure

Verbetering van epidemiologisch onderzoek naar gezondheidseffecten van blootstelling aan radiofrequente elektromagnetische velden

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 12 april 2022
des middags te 2.15 uur

door

Marije Reedijk

geboren op 11 maart 1988

te Utrecht

Promotoren:

Prof. dr. ir. R.C.H. Vermeulen

Prof. dr. ir. H. Kromhout

The studies in this thesis were financially supported by:

- ZonMW, the Neterlands Organisation for Health Research within the research program Electromagnetic Fields and Health Research under grant numbers 85200001, 85200002, 85800001.
- The GERoNiMO project is supported by the European Union (grant 603794).

TABLE OF CONTENTS

Chapter 1

General introduction

Currently everybody is exposed to radiofrequency electromagnetic fields (RF-EMF) in daily life, either by their own communications device(s) or by mobile phone base stations or radio transmitters in their living environment. The amount of sources emitting RF-EMF has rapidly increased during the past decade mainly as a result of the introduction of wireless information technologies such as smartphones, tablets and Wi-Fi. For environmental exposures, such as RF-EMF, even a small increase in risk may translate into a large burden of disease in the general population since everybody is in a greater or lesser extent continuously exposed. Despite that much experimental and epidemiological research already has been conducted on potential adverse health effects of RF-EMF there still remain concerns about these effects both in society as among scientists[1–4].

Several authoritative reports have been previously written on the effects of RF-EMF exposure. In 2011, the International Agency for Research on Cancer (IARC) has classified RF-EMF as possible carcinogenic (2B) to humans[2,5]. The UK Health Protection Agency (nowadays Public Health England) reported that the results for health effects of RF-EMF exposure are still inconclusive as there is no evidence for an effect but also no strong evidence against it[4]. The concerns in society were intensified again by the recent introduction and roll-out of the fifth generation communication network (5G) in the Netherlands. As a result the Health Council of the Netherlands (Dutch: Gezondheidsraad) was asked recently by the House of Representatives of the Dutch parliament to give an advise on potential health effects of 5G EMF exposure[3]. Based on the literature, including both experimental and epidemiological studies, the Dutch Health Council concluded that there are indications that electromagnetic fields with the frequencies of 5G have the potential to harm a range of health outcomes, including cancer, male infertility and birth outcomes and defects, while uncertainties remain for other diseases like neurological outcomes[3].

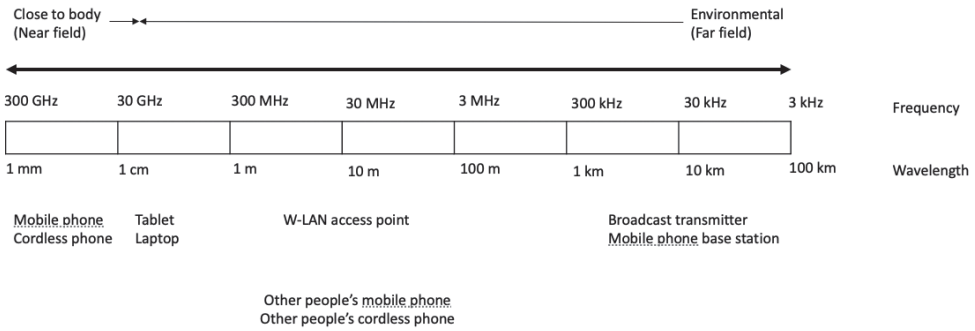
There are still uncertainties in RF-EMF epidemiology including the scarcity of longitudinal studies in the field which makes it hard to investigate causality, the lack of a clear biological mechanism and the lack of a reliable exposure assessment methods of RF-EMF that can be applied in large observational epidemiological studies. Exposure assessment of RF-EMF is challenging as electromagnetic fields are spatially and temporally variable and a lot of factors play a role including people's behaviour and perception. This thesis addresses some of the uncertainties in RF-EMF epidemiology such as investigating the repeatability of self-reported use and improving exposure assessment of mobile phone use by including both self-reported and operator-recorded data. Furthermore the role of perceived exposure next to the actual, modeled, exposure will be investigated.

RF-EMF

Radiofrequency electromagnetic fields describes the part of the spectrum of electromagnetic radiation with frequencies ranging from 3 kilohertz to 300 gigahertz (Figure 1). RF-EMF are emitted by both

natural and artificial sources. Natural sources contribute little to the total exposure of humans to RF-EMF[6]. RF-EMF sources can be further categorized by the distance of the source to the human body, which is divided in near- and far-field sources (Figure 1)[6]. Mobile and cordless phones are examples of near-field sources whereas mobile phone base stations and broadcast transmitters are examples of far-field sources. Near-field sources, typically less than one wavelength away from a transmitter, are very localized (e.g. head, hand), while for far-field exposure sources, more than one wavelength away, the whole body is exposed[6]. For most persons the highest RF-EMF exposure levels for the whole brain will likely be caused by (using) mobile phones, however the relative importance of other sources, such as far-field sources as mobile phone base stations, will differ per organ-system of interest and underlying study population[7,8]. For example higher overall doses were found in younger age groups for an integrated exposure model of RF-EMF exposure, developed by van Wel et al., including both near- and far fields sources such as mobile phone use (2G/3G), cordless phone, tablet, laptop and Wi-Fi[8]. Also differences in near- and far-field exposure were shown between study populations in France and Spain[8]. Many factors influence RF-EMF exposure levels such as the different near- and far field sources, time using the source or spent in a location near a source, active or passive use of the device, transmission output, position of the device or transmitter and personal factors such as age[9].

Figure 1. Near and far field RF-EMF sources, based on[6]



Exposure assessment of RF-EMF

Almost a decade ago the UK Health Protection Agency argued that improving exposure assessment of RF-EMF for epidemiological studies is essential to produce more conclusive answers on the question if RF-EMF exposure is harmful for humans[4,10]. Since then improvements have been made in the exposure assessment of RF-EMF including the usage of data collected by mobile phone providers and smartphone applications. In this thesis we focused on RF-EMF exposure from mobile phones and mobile phone base stations as the relative contribution of these sources are still the biggest contributors for respectively the whole-brain and the whole-body[8].

RF-EMF exposure assessment from mobile phones

Several methods exist to take RF-EMF exposure from mobile phone use into account in epidemiological investigations: self-reported information on use, information collected by providers, and/or data retrieved from smartphone applications like XMobisense[11,12]. Epidemiological investigations into possible health effects of mobile phones still often use self-reported mobile phone use via questionnaires or interviews for exposure assessment with a high potential for recall bias and imprecision[13–16]. These errors can affect health risk estimates as either under- or overestimation occurs depending on different factors e.g. age, amount of phone use (both frequency and number) and possibly disease status, which has been shown in several studies in the past[13–16]. However little is known about how well participants can repeat their historical mobile phone use by self-reported information.

The second method, retrieving operator-recorded data from mobile phone providers would in principle be a more objective method to assess RF-EMF exposure from mobile phone use. These data are often retrieved from billing systems, which make them less prone to recall bias. However due to issues with subscription in employer's names, prepaid services, no access to provider data, attrition of participants over time due to change of provider or lacking informed consent it is hard to receive operator data for all participants in large (prospective) epidemiological investigations.

More recently, through the introduction of smartphones, another possibility was introduced by installing monitoring applications onto the phone. Although feasibility and reliability has been shown in smaller studies, the applicability of this method to larger study populations is difficult and the collected information will not provide insight into long-term exposures[11,12].

It could be very beneficial to explore statistical methods that allow for a combination of more than one data source, resulting in a more complete, valid and precise estimate of mobile phone use over a relevant time period. An example of such an approach is regression calibration, which has been applied for text messages by Redmayne et al[17] and are increasingly used in environmental epidemiology[18–21].

RF-EMF exposure assessment from mobile phone base stations

Exposure assessment of RF-EMF from mobile phone base stations is complex because of the large spatial variation in exposure levels and various variables that play a role (e.g. building height, building characteristics, antenna directions)[22]. In the past several methods have been developed to assess exposure to RF-EMF from mobile phone base stations e.g. measurements, spatial interpolation, and distance to the nearest antenna[6]. In this thesis we used NISMap, a geospatial radio wave propagation model, developed by Bürgi et al and improved by Beekhuizen et al for the Dutch situation[23–26]. NISMap has been proven to work sufficiently in epidemiological studies, with a spearman correlation of 0.48 between model predictions and 48 hour measurements[27]. A study by Martens et al. showed moderate

Spearman's correlations between modeled predictions and measurements at home (R_{Sp}=0.51), measurements in the bedroom (R_{Sp}=0.41), and over a 24 h period (R_{Sp}=0.36)[28]. Many of the current exposure assessment tools for RF-EMF exposure, such as NISMap, will become largely inadequate for the 5G environment, which is fundamentally different in exposure compared to previous technologies (2G-4G networks). The research in this thesis was conducted before the recent introduction of the 5G technology.

Psychosocial mechanisms

Experimental studies have shown that not only the physical component of exposures can influence health but also psychological factors such as perception, worries or concerns can play a role in the reporting of health complaints[29–31]. There are suggestions that individuals may attribute symptoms to environmental exposures by negative expectations, which has been described as the nocebo-effect, the counterpart of the more known placebo-effect[32–35]. For RF-EMF exposure from mobile phone base stations several studies have shown stronger associations with health complaints for the perceived exposure than with the actual, modeled, exposure levels[32,33,35–37]. Furthermore evidence on direct or indirect effects of perceived exposures is currently lacking as only a few longitudinal studies have been performed and many studies used (only) self-reported data[38–40]. Within environmental epidemiology the influence of perceptions and modern health worries on developing acute and chronic disease, either direct or indirect, is still a largely neglected topic for some exposures[41]. Therefore is important to explore if these mechanism found for RF-EMF from mobile phone base stations can also be found for other environmental exposures like road traffic noise and air pollutants. For psychosocial factors other intervention strategies, such as risk communication strategies, are necessary to prevent symptom reporting then for the effects of the exposure itself.

Combining multiple studies

Health effects of environmental exposure, including RF-EMF, are generally hard to detect because effects are relatively small in magnitude and other risk factors may have stronger effects than the exposures under study[42,43]. To gain enough statistical power to study rare health outcomes of interest, like brain cancer or neurological diseases for RF-EMF exposure, it is necessary to increase the sample size of the study population. This can be done by combining (health and exposure) data from multiple epidemiological studies. There are several ways to accomplish data sharing, which might be driven by ethical, legal and practical reasons. Data from epidemiological studies can be pooled by a harmonized and standardized protocol beforehand or by retrospectively harmonized and standardized data. Furthermore data can be analysed on an individual level, aggregated level or through meta-analyses on an individual or study level. Combining data on an individual level could result in more statistical

power and offers more flexibility in the data analysis[44]. In this thesis we have used several methods to combine data from multiple (prospective) cohort studies.

Assessment of health

For human observational studies it is not only important to accurately measure the exposure of interest, in this thesis RF-EMF exposure, but also the health outcome of interest[45]. Research instruments that can be used for heath outcome assessment can be either self-reported (e.g. questionnaires, interview) or more objective information (e.g. medical records, laboratory results, physical measurements, registry data). The used instruments for both the exposure and health outcome should preferable be valid, precise and reproducible to accurately measure what is intended to be measured, which should be consistent over time and among different groups and studies[45]. Extensive information on health is captured in health registries which can be linked (through personal information) with epidemiological studies. These registries capture pertinent information on disease aspects, for example for cancer outcomes information is collected about diagnose date, cancer classification and information on metastasis amongst others. Disease standardization is achieved by classification systems, such as the International Classification of Diseases (ICD-10), which is currently widely used in the health care systems[46]. In the Netherlands for diseases as cancer and cause-specific mortality well developed registries are in place[47,48]. However this is not the case for all diseases for example for neurological diseases no specific registry is currently available in the Netherlands so information on disease status needs to be obtained from other information sources. For neurological diseases, such as Amyotrophic Lateral Sclerosis and Parkinson's Disease there is limited experimental and epidemiological evidence for RF-EMF exposure[3].

AIM THESIS

In this thesis I have addressed several important uncertainties in human observational studies on health effects of RF-EMF. We improved exposure assessment of RF-EMF by including both self-reported and operator-recoded data in Cohort Study of Mobile Phone Use and Health (COSMOS), we developed a case-ascertainment method for Parkinson's Disease, and examined the role of perceived exposure in combination with the role of actual, modelled, exposure. *These investigations will be instrumental to expand the knowledge on possible associations between RF-EMF exposure from mobile phones and mobile phone base stations and health, the overall aim of my PhD project.* The first part of this thesis is about improving RF-EMF epidemiology by combining data from multiple prospective cohort studies and by evaluating and determining exposure assessment of RF-EMF by investigating both the validity and reliability of mobile phone use. In the second part of this thesis we developed a method to identify Parkinson's Disease in ongoing prospective cohorts. In the last part of this thesis we expanded our

knowledge on possible health effects of RF-EMF by investigating the association between maternal mobile phone use and birth outcomes in four birth cohorts. Furthermore we investigated the role of perceived exposure next to the modeled exposure on self-reported health outcomes for RF-EMF from mobile phone base stations, road traffic noise and several air pollutants.

THESIS OUTLINE

In **chapter two** the establishment of and the data collected within the prospective cohort study LIFEWORK, the Dutch contribution to the international prospective COSMOS cohort study, is described. This chapter includes a methodological study on the repeatability of the questionnaire including questions on mobile phone use. In **chapter three** we evaluated and compared several statistical approaches, including regression calibration, to combine self-reported and operator-recorded mobile phone use data within the international COSMOS study among more than 290.000 participants in five European countries[49,50]. **Chapter four** describes the case ascertainment algorithm we developed to identify participants with Parkinson's Disease within two prospective cohorts: AMIGO (Occupational and Environmental Health Cohort study) and EPIC-NL (European Prospective Investigation into Cancer and Nutrition in the Netherlands). For this algorithm information from multiple data sources (self-reported and registry-based) are combined. Subsequently, the algorithm was validated with evidence retrieved from general practitioners. Associations with known etiological factors (smoking, family history and sex) were conducted to establish internal face-validity. **Chapter five** provides further insights into the associations between self-reported maternal mobile phone use during pregnancy and several birth outcomes, including fetal growth, pregnancy duration and birth weight. We investigated this in four birth cohorts[51–54]. In **chapter six** we evaluated if earlier findings on perceived RF-EMF exposure from mobile phone base stations[33,35] were also applicable for other environmental exposures such as air pollutants and road traffic noise. Therefore we examined the associations between RF-EMF from base stations, road traffic noise and air pollutants on self-reported health (non-specific symptoms, sleep quality, respiratory symptoms) for both modeled and perceived exposure within the AMIGO cohort. Finally, in **chapter seven**, I will discuss the main findings of this thesis, which I will place in a broader context within the RF-EMF research field and describe opportunities and challenges of using multiple information sources.

Cohorts used

In this dissertation we used data from different ongoing epidemiological investigations in the field of RF-EMF, which are briefly described below.

LIFEWORK

In 2011 the nationwide prospective cohort LIFEWORK was established to study occupational and environmental health. LIFEWORK is the Dutch contribution to the international COSMOS study (described below) and has a specific focus on EMF exposure from mobile phones. The establishment of and collected data within LIFEWORK is described in more detail in the second chapter of this thesis. For LIFEWORK three large prospective Dutch cohorts collaborated: EPIC-NL[55], Nightingale[56] and AMIGO[57]. We preformed case-ascertainment of Parkinson's Disease between 2015 and 2017 for participants in the cohorts EPIC-NL and AMIGO (chapter four).

COSMOS

With nearly 90.000 participants LIFEWORK is the second largest contributor to the international COSMOS study which aim is to evaluate health effects of RF-EMF from mobile communication devices. COSMOS is an international consortium of Denmark, Finland, Sweden, the Netherlands, UK and France. Data from these countries will be combined to have sufficient power to study mobile phone use in relation to rare diseases nearby the head such as brain cancer and neurological disorders including Parkinson's disease. More details on the COSMOS study are described in previous publications[49,58–60]. An important asset of COSMOS is the prospective design, reducing recall and selection bias which is a limitation of previous large case-control studies, such as INTERPHONE[61], on mobile phone use and health. Data from the international COSMOS study is used in chapter three of this thesis.

GERoNiMO

For the research conducted in chapter five data is used from four birth cohorts namely: the Amsterdam Born Children and their Development Study (ABCD)[52], the Danish National Birth Cohort (DNBC)[51], the Spanish Environment and Childhood Project (INMA)[62], and the Mothers and Children's Environmental Health (MOCEH) study[54]. These data were pooled for the GERoNiMO project, an European project on Generalized EMF research using novel methods. The aim of the GERONIMO project was to resolve knowledge gaps on EMF and health and to propose non-technological methods to reduce EMF exposure[63].

REFERENCES

1. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Non-ionizing radiation, Part 2: Radiofrequency electromagnetic fields. *IARC Monogr. Eval. Carcinog. Risks Hum.* **102**, 1–460 (2013).
2. Baan, R. *et al.* Carcinogenicity of radiofrequency electromagnetic fields. *Lancet Oncol.* **12**, 624–626 (2011).
3. Gezondheidsraad. advies 5G en gezondheid. 2–34 (2020).
4. HPA. Health Effects from Radiofrequency Electrmagnetic Fields. Independent Advisory Group on Non-Ionising Radiotion. www.hpa.org.uk (2012).
5. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Non-ionizing radiation, Part 2: Radiofrequency electromagnetic fields. *IARC Monogr. Eval. Carcinog. Risks Hum.* **102**, 1–460 (2013).
6. Frei, P. & Röösli, M. Exposure to radiofrequency electromagnetic fields in our everyday environment. in *Epidemiology of Electromagnetic Fields* 125–140 (2014). doi:10.1201/b16968.
7. Bolte, J. F. B. & Eikelboom, T. Personal radiofrequency electromagnetic field measurements in the Netherlands: Exposure level and variability for everyday activities, times of day and types of area. *Environ. Int.* **48**, 133–142 (2012).
8. van Wel, L. *et al.* Radio-frequency electromagnetic field exposure and contribution of sources in the general population: an organ-specific integrative exposure assessment. *J. Expo. Sci. Environ. Epidemiol.* (2021) doi:10.1038/s41370-021-00287-8.
9. van Wel, L. *Challenges and advancements in radiofrequency electromagnetic fields exposure assessment for environmental epidemiology*. (2020).
10. Toledano, M. B. & Smith, R. B. Future of EMF epidemiology. *Epidemiol. Electromagn. Fields* 327–342 (2014) doi:10.1201/b16968.
11. Goedhart, G., Kromhout, H., Wiart, J. & Vermeulen, R. Validating self-reported mobile phone use in adults using a newly developed smartphone application. *Occup. Environ. Med.* **72**, oemed-2015-102808 (2015).
12. Goedhart, G. *et al.* Recall of mobile phone usage and laterality in young people: The multinational Mobi-Expo study. *Environ. Res.* **165**, 150–157 (2018).
13. Vrijheid, M., Deltour, I., Krewski, D., Sanchez, M. & Cardis, E. The effects of recall errors and of selection bias in epidemiologic studies of mobile phone use and cancer risk. *J. Expo. Sci. Environ. Epidemiol.* **16**, 371–84 (2006).
14. Vrijheid, M. *et al.* Recall bias in the assessment of exposure to mobile phones. *J. Expo. Sci. Environ. Epidemiol.* **19**, 369–381 (2009).
15. Aydin, D. *et al.* Impact of random and systematic recall errors and selection bias in case-control studies on mobile phone use and brain tumors in adolescents (CEFALO study). *Bioelectromagnetics* **32**, 396–407 (2011).
16. Roosli, M. *Epidemiology of electromagnetic fields* . (2014).
17. Redmayne, M., Smith, E. & Abramson, M. J. A forecasting method to reduce estimation bias in self-reported cell phone data. *J. Expo. Sci. Environ. Epidemiol.* **23**, 539–544 (2013).
18. Bateson, T. F. & Wright, J. M. Regression calibration for classical exposure measurement error in environmental epidemiology studies using multiple local surrogate exposures. *Am. J. Epidemiol.* **172**, 344–352 (2010).
19. Fraser, G. E. & Stram, D. O. Regression calibration in studies with correlated variables measured with error. *Am. J. Epidemiol.* **154**, 836–844 (2001).
20. Roosbroeck, S. Van *et al.* Traffic-Related Outdoor Air Pollution and Respiratory Symptoms in Children: The Impact of Adjustment for Exposure Measurement Error. *Epidemiology* **19**, 409–416 (2008).
21. Petit, C. *et al.* Association of environmental insecticide exposure and fetal growth with a bayesian model including multiple exposure sources: The PELAGIE mother-child cohort. *Am. J. Epidemiol.* **175**, 1182–1190 (2012).

22. Beekhuizen, J., Kromhout, H., Bürgi, A., Huss, A. & Vermeulen, R. What input data are needed to accurately model electromagnetic fields from mobile phone base stations? *J. Expo. Sci. Environ. Epidemiol.* **25**, 53–57 (2015).

23. Beekhuizen, J. *et al.* Modelling indoor electromagnetic fields (EMF) from mobile phone base stations for epidemiological studies. *Environ. Int.* **67**, 22–6 (2014).

24. Bürgi, A., Theis, G., Siegenthaler, A. & Röösli, M. Exposure modeling of high-frequency electromagnetic fields. *J. Expo. Sci. Environ. Epidemiol.* **18**, 183–191 (2008).

25. Bürgi, A. *et al.* A model for radiofrequency electromagnetic field predictions at outdoor and indoor locations in the context of epidemiological research. *Bioelectromagnetics* **31**, 226–236 (2010).

26. Beekhuizen, J., Vermeulen, R., Kromhout, H., Bürgi, A. & Huss, A. Geospatial modelling of electromagnetic fields from mobile phone base stations. *Sci. Total Environ.* **445–446**, 202–209 (2013).

27. Martens, A. L. *et al.* Residential exposure to RF-EMF from mobile phone base stations: Model predictions versus personal and home measurements. *Sci. Total Environ.* **550**, 987–993 (2016).

28. Martens, A. L. *et al.* Validity of at home model predictions as a proxy for personal exposure to radiofrequency electromagnetic fields from mobile phone base stations. *Environ. Res.* **142**, 221–226 (2015).

29. Szemerszky, R., Köteles, F., Lihi, R. & Bárdos, G. Polluted places or polluted minds? An experimental sham-exposure study on background psychological factors of symptom formation in 'Idiophatic Environmental Intolerance attributed to electromagnetic fields'. *Int. J. Hyg. Environ. Health* **213**, 387–94 (2010).

30. Crichton, F., Dodd, G., Schmid, G., Gamble, G. & Petrie, K. J. Can expectations produce symptoms from infrasound associated with wind turbines? *Health Psychol.* **33**, 360–4 (2014).

31. Witthöft, M. & Rubin, G. J. Are media warnings about the adverse health effects of modern life self-fulfilling? An experimental study on idiopathic environmental intolerance attributed to electromagnetic fields (IEI-EMF). *J. Psychosom. Res.* **74**, 206–212 (2013).

32. Baliatsas, C. *et al.* Non-specific physical symptoms and electromagnetic field exposure in the general population: Can we get more specific? A systematic review. *Environ. Int.* **41**, 15–28 (2012).

33. Martens, A. L. *et al.* Modeled and Perceived Exposure to Radiofrequency Electromagnetic Fields From Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort. *Am. J. Epidemiol.* **186**, 210–219 (2017).

34. Hahn, R. A. The Nocebo Phenomenon: Concept, Evidence, and Implications for Public Health. *Prev. Med. (Baltim).* **26**, 607–611 (1997).

35. Baliatsas, C. *et al.* Actual and perceived exposure to electromagnetic fields and non-specific physical symptoms: An epidemiological study based on self-reported data and electronic medical records. *Int. J. Hyg. Environ. Health* **218**, 331–344 (2015).

36. Fyhri, A. & Klæboe, R. Road traffic noise, sensitivity, annoyance and self-reported health-A structural equation model exercise. *Environ. Int.* **35**, 91–97 (2009).

37. Héritier, H. *et al.* The Association between Road Traffic Noise Exposure, Annoyance and Health-Related Quality of Life (HRQOL). *Int. J. Environ. Res. Public Health* **11**, 12652–12667 (2014).

38. Baliatsas, C., van Kamp, I., Hooiveld, M., Lebret, E. & Yzermans, J. The relationship of modern health worries to non-specific physical symptoms and perceived environmental sensitivity: A study combining self-reported and general practice data. *J. Psychosom. Res.* **79**, 355–361 (2015).

39. Köteles, F., Szemerszky, R., Freyler, A. & Bárdos, G. Somatosensory amplification as a possible source of subjective symptoms behind modern health worries. *Scand. J. Psychol.* **52**, 174–178 (2011).

40. Petrie, K. J. *et al.* Worries about modernity predict symptom complaints after environmental pesticide spraying. *Psychosom. Med.* **67**, 778–782 (2005).

41. Cori, L., Donzelli, G., Gorini, F., Bianchi, F. & Curzio, O. Risk perception of air pollution: A systematic review focused on particulate matter exposure. *Int. J. Environ. Res. Public Health* **17**, 1–27 (2020).
42. Beaker, D. & Nieuwenhuijsen, M. J. *Environmental epidemiology, study methods and applications*. (2008).
43. Pekkanen, J. & Pearce, N. Environmental Epidemiology : Challenges and Opportunities. *Ehp* **109**, 1–5 (2001).
44. Gaye, A. *et al.* DataSHIELD: Taking the analysis to the data, not the data to the analysis. *Int. J. Epidemiol.* **43**, 1929–1944 (2014).
45. Nieuwenhuijsen, M. J. *Exposure Assessment in Environmental Epidemiology*. (2015).
46. WHO. *ICD-10: international statitical classification of diseases and related health problems*. (2004).
47. CBS. Cause of death register. *CBS website* (2016).
48. IKNL. https://iknl.nl/
49. Schüz, J. *et al.* An international prospective cohort study of mobile phone users and health (Cosmos): design considerations and enrolment. *Cancer Epidemiol.* **35**, 37–43 (2011).
50. Toledano, M. B., Smith, R. B., Chang, I., Douglass, M. & Elliott, P. Cohort profile: UK COSMOS—a UK cohort for study of environment and health. *Int. J. Epidemiol.* dyv203 (2015) doi:10.1093/ije/dyv203.
51. Olsen, J. *et al. The Danish National Birth Cohort ± its background, structure and aim*.
52. Van Eijsden, M., Vrijkotte, T. G. M., Gemke, R. J. B. J. & Van der Wal, M. F. Cohort profile: The Amsterdam born children and their development (ABCD) study. *International Journal of Epidemiology* vol. 40 1176–1186 (2011).
53. Guxens, M. *et al.* Cohort profile: The INMA-INfancia y Medio Ambiente-(environment and childhood) project. *Int. J. Epidemiol.* **41**, 930–940 (2012).
54. Kim, B. M. *et al.* The Mothers and Children's Environmental Health (MOCEH) study. *Eur. J. Epidemiol.* **24**, 573–583 (2009).
55. Beulens, J. W. J. *et al.* Cohort Profile: The EPIC-NL study. *Int. J. Epidemiol.* **39**, 1170–1178 (2010).
56. Pijpe, A. *et al.* The Nightingale study: rationale, study design and baseline characteristics of a prospective cohort study on shift work and breast cancer risk among nurses. *BMC Cancer* **14**, 47 (2014).
57. Slottje, P., Yzermans, C. J., Korevaar, J. C., Hooiveld, M. & Vermeulen, R. C. H. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in The Netherlands. *BMJ Open* **4**, e005858 (2014).
58. Toledano, M. B., Smith, R. B., Brook, J. P., Douglass, M. & Elliott, P. How to establish and follow up a large prospective cohort study in the 21st century - Lessons from UK COSMOS. *PLoS One* **10**, 1–16 (2015).
59. Toledano, M. B., Smith, R. B., Brook, J. P., Douglass, M. & Elliott, P. How to Establish and Follow up a Large Prospective Cohort Study in the 21st Century - Lessons from UK COSMOS. *PLoS One* **10**, e0131521 (2015).
60. Heinävaara, S., Tokola, K., Kurttio, P. & Auvinen, A. Validation of exposure assessment and assessment of recruitment methods for a prospective cohort study of mobile phone users (COSMOS) in Finland: a pilot study. *Environ. Heal.* **10**, 14 (2011).
61. Cardis, E. Brain tumour risk in relation to mobile telephone use: Results of the INTERPHONE international case-control study. *Int. J. Epidemiol.* **39**, 675–694 (2010).
62. Guxens, M. *et al.* Cohort profile: The INMA-INfancia y Medio Ambiente-(environment and childhood) project. *Int. J. Epidemiol.* **41**, 930–940 (2012).
63. Geronimo project. https://www.isglobal.org/en/-/geronimo-generalized-emf-research-using-novel-methods-an-integrated-approach-from-research-to-risk-assessment-and-support-to-risk-management.

Chapter 2

Cohort profile: LIFEWORK, a prospective cohort study on occupational and environmental risk factors and health in the Netherlands

Marije Reedijk[a], Virissa Lenters[a], Pauline Slottje[b], Anouk Pijpe[b], Petra H. Peeters, Joke C. Korevaar, Bas Bueno-de-Mesquita, W.M. Monique Verschuren, Robert A. Verheij, Inka Pieterson, Flora E. van Leeuwen, Matti A. Rookus[c], Hans Kromhout[c], Roel Vermeulen[c]

[a] Equal contribution
[b] Equal contribution
[c] Equal contribution

Published: BMJ Open. 2018 ; Vol. 8, No. 2.

ABSTRACT

LIFEWORK is a large federated prospective cohort established in the Netherlands to quantify the health effects of occupational and environmental exposures. This cohort is also the Dutch contribution to the international COSMOS study on mobile phone use and health. In this paper we describe the study design, ongoing data collection, baseline characteristics of participants, and the repeatability of key questionnaire items. 88 466 participants were enrolled in three cohort studies in 2011–2012. Exposure information was collected by a harmonized core questionnaire, or modeled based on occupational and residential histories; domains include air pollution (e.g., NO_2, $PM_{2.5}$), noise, electromagnetic fields (EMF), mobile phone use, shift work, and occupational chemical exposures. Chronic and sub-acute health outcomes are assessed by self-report and through linkage with health registries. Participants had a median age of 51 years at baseline (range 19–87), and the majority is female (90%), with nurses being overrepresented. Median exposure levels of NO_2, $PM_{2.5}$, EMF from base stations, and noise at the participants' home addresses at baseline was 22.9 µg/m^3, 16.6 µg/m^3, 0.003 mWm^2, and 53.1 dB, respectively. Twenty-two percent of participants reported to have started using a mobile phone more than 10 years prior to baseline. Repeatability for self-reported exposures was moderate to high (weighted kappa range: 0.69–1) for a subset of participants (n=237) who completed the questionnaire twice. We are actively and passively observing participants; we plan to administer a follow-up questionnaire every 4-5 years—the first follow-up was completed in 2017—and linkage to cause-of-death and cancer registries occurs on a (bi)annual basis. This prospective cohort offers a unique, large and rich resource for research on contemporary occupational and environmental health risks and will contribute to the large international COSMOS study on mobile phone use and health.

INTRODUCTION

People are exposed to a myriad of occupational and environmental circumstances throughout their lifetime. Some exposures, such as air pollution, are well-established risk factors for acute and chronic disease outcomes; for others, like electromagnetic fields (EMF), uncertainty remains about their effects on health[1,2]. Until now the majority of research has evaluated health effects of single exposures, while occupational and environmental exposures often occur in complex mixtures (either simultaneous and/or intermittent) and at low concentrations[3,4]. Therefore, to obtain more accurate exposure–health effect estimates, it is important to assess the effects of multiple simultaneous exposures on different health outcomes during the entire life course by using state-of-the-art exposure assessment tools and information[5]. This has recently been formalized in the exposome concept, which proposes to study the totality of environmental exposures over the life course[6].

Although the health effects of environmental exposures are frequently small in magnitude, the disease burden is potentially large, as the prevalence of many exposures is high. The Global Burden of Disease study estimated that about half of global deaths are attributable to modifiable exposures such as particulate air pollution (both household and ambient), smoking, physical activity and diet[7]. Nevertheless, a large proportion of disease caused by modifiable factors remains unexplained[8,9]. Although a couple of large occupational and environmental cohorts have recently been established, such as CONSTANCES[10], occupational and environmental risk factors for health have generally been studied either in narrowly focused, mainly retrospective studies, or more broadly as an additional topic in ongoing cohort studies which initially had another focus.

Therefore, we established a large prospective cohort study, LIFEWORK, comprising three cohorts, with a focus on comprehensively assessing occupational and environmental exposures, and health[11-13]. We gathered information on participants' complete residential and occupational histories in an effort to capture changes in important aspects of the external exposome throughout the life course, and also gathered information on (sub)acute and chronic health outcome data. This cohort also contributes to the six-country effort to investigate the potential health effects of long-term use of mobile phones (the COhort Study of MObile phone uSe and health (COSMOS) study[14]). Both self-reported as well as operator-recorded mobile phone use data is collected over time, allowing for improved exposure assessment[15]. Another area of interest is how health is affected by communication technologies (e.g., blue light) and shift work, which drive the modern 24-hr economy.

Our study was designed to assess a broad array of environmental and occupational exposures, with a focus on mobile phone use and exposure to electromagnetic fields. In this paper, we present the LIFEWORK cohort by describing the study design, data collection, baseline characteristics, exposure distributions, the repeatability of several key questionnaire items, and projected study power for several outcomes.

COHORT DESCRIPTION

Study design and procedures

The prospective LIFEWORK cohort was established to collect a large amount of high quality data on occupational and environmental exposures using a harmonized core questionnaire. To achieve a large study population, data collection was integrated into an already existing cohort (EPIC-NL, the European Prospective Investigation into Cancer and Nutrition in the Netherlands) and two *de novo* cohorts (the Nightingale Study and AMIGO, the Occupational and Environmental Health Cohort Study). From the outset, the LIFEWORK cohort was designed as a federated study, whereby the subcohorts are governed by an overarching governance board alongside independent governance boards, and data collection was designed to capture a core set of exposures and outcomes of interest. In addition to the harmonized core data collected on exposures and health outcomes, each of these cohorts collected data to fulfill additional research aims (described below in the data collection section). Deviations in data collection across the three subcohorts are briefly described below and are detailed in Supplementary Table S1, including which data required harmonization.

Study designs and recruitment for the individual cohorts have previously been described[11-13]. Briefly, the LIFEWORK cohort comprises (1) female (current and former) nurses, aged 18-65 years at enrollment in 2011, who were recruited via the registration system for healthcare professionals, i.e. the BIG-register (the Nightingale Study)[13]; (2) adults from the general population, aged 31-65 years at recruitment in 2011-2012, who were recruited, with a maximum of one person per household, via a national general practitioners network, i.e. the NIVEL Primary Care Database (AMIGO)[12]; and (3) participants from the EPIC-NL cohort. EPIC-NL participants were originally recruited between 1993–1997 (n=40 011), either via a breast cancer screening program conducted in Utrecht and neighboring towns (women aged 49-70 years; EPIC-NL Prospect), or from the general population of three cities, Amsterdam, Doetinchem, and Maastricht (adults aged 20-59 years; EPIC-NL MORGEN; participants from Doetinchem were not invited for LIFEWORK)[11]. The LIFEWORK baseline questionnaire was the baseline questionnaire for AMIGO and Nightingale (2011-2012) and the third questionnaire for EPIC-NL (2011). The studies were approved by local ethical review committees, and participants signed an informed consent form for each subcohort prior to enrolment.

88 731 participants (of whom 88 466 are unique) were enrolled in 2011 and 2012 upon completion of a self-administered paper questionnaire (EPIC-NL) or web-based questionnaire (with the option to receive a paper version; Nightingale and AMIGO). With 59 941 participants (67%), the Nightingale Study is the largest cohort contributing to LIFEWORK, followed by AMIGO (n=14 829, 17%) and EPIC-NL (n=13 961, 16%). The participation rates were 16% for AMIGO, 31% for Nightingale, and 51% for EPIC-NL (response in 2011-2012, the third follow-up; and 40% at EPIC-NL baseline in 1993-1997). Compared

to the 40 011 who enrolled at EPIC-NL baseline (1993-1997), the 13 961 EPIC-NL study participants who participated in LIFEWORK (i.e., responded to the third EPIC-NL questionnaire in 2011/2012) had a similar age distribution (mean age of 50 vs. 49 years), and a greater proportion were higher educated (28 % vs. 20%) and were women (80% vs. 74%).

Participants were asked to consent to linkage to the Municipal Personal Records Database and health registries (e.g., cancer and mortality). Consent was a prerequisite for participation in AMIGO and Nightingale and was therefore 100%. In EPIC-NL, consent for linkage was obtained for 97% at baseline (1993-1997)[11]. Participants were also asked at enrollment (2011-2012) to consent to linkage to their mobile phone operators to obtain data on their monthly mobile phone use, although it was clearly stated that this was optional and not a prerequisite for participation. 35 966 participants (41%) consented to this.

Duplicate enrollments

Participants were independently invited to participate in the three cohort studies, and the source populations overlapped somewhat. It was therefore possible that participants enrolled in more than one subcohort, and received the LIFEWORK baseline questionnaire more than once. 265 participants completed more than one LIFEWORK questionnaire (15 both AMIGO and EPIC-NL, 184 both AMIGO and Nightingale, and 66 both Nightingale and EPIC-NL).

Data collection

Questionnaire data

Topics included in the LIFEWORK baseline questionnaire are listed in Table 1 and elaborated in Supplementary Table S1. Information on mobile phone use and other wireless technologies was assessed according to the COSMOS study protocol[14]. Validated questionnaires were used to assess various health outcomes, many of which were also included in COSMOS[16-20]. Occupational EMF exposures were assessed considering sources (e.g., MRI equipment and dielectric heating) with high EMF exposure levels. Information on an extensive set of potential confounders, including socio-demographic, socio-economic, and lifestyle factors was collected, along with information on important modifiers of exposure and co-exposures.

Table 1. Key topics in the LIFEWORK baseline questionnaire and the models used to assess environmental and occupational exposures.

Questionnaire	
Topic	**Items**
Characteristics	Age, sex, marital status, country of birth, height, weight
Socio-economic	Employment status, education
Lifestyle	Smoking, alcohol use, physical activity[a]
Housing characteristics	Bedroom floor, position of the bedroom relative to the street
Residential history	Full residential history of addresses lived at for at least 12 months
Occupational history	AMIGO/EPIC-NL: Full history of jobs performed for at least 6 months (job title, type of company, average number of hours per week), shift work (ever shift work and if permanent or rotating shift system)[a]
	The Nightingale Study: Full history of jobs performed for at least 6 months (job title, type of company, average number of hours per week, shiftwork, physical activity), shift work (rotation, frequency, calendar years), chronotype (MEQ)[22a]
EMF exposure	Current and historical mobile phone use, cordless phone use and internet mobile phone use[b], job tasks or use of equipment likely to lead to high EMF exposure
Health	*Current*: physical and mental well-being (SF-12), headache (HIT-6)[16], migraine (ID-migraine)[17], sleep quality (MOS sleep scale)[18], memory problems, hearing problems, tinnitus *Past*: doctor assessed symptoms and diseases, and age at diagnosis for conditions and events including diabetes, cardiovascular events, neurodegenerative diseases, cancer, asthma[a]
	Reproductive health (women): parity, birth outcomes, hormone use[a]
Modeled exposures	
Geospatial model	**Environmental exposures**
NISMap[23,24]	RF-EMF from mobile phone base stations, TV and radio antennas
ESCAPE LUR-model[25,26]	Air pollutants ($PM_{2.5}$, $PM_{2.5}$ absorbance, PM_{10}, NO_2, OP_{dtt})
NDVI[27]	Green space measure
STAMINA[28]	Traffic noise exposure (e.g., L_{den})
LOCATUS[29]	Built environment (e.g., number of fast-food restaurants and sport facilities in surrounding area)
Job-exposure matrix	**Occupational exposures**
ALOHA+ JEM[30,31]	Dust (biological, mineral dust), pesticides (all pesticides, herbicides, insecticides, fungicides), solvents (total solvents, aromatic solvents, chlorinated solvents)
DOM-JEM[32,33]	Diesel motor exhaust fumes, asbestos, chromium, nickel, PAHs, silica, animal dust, biological dust, endotoxin
Shock-JEM[34]	Electric shocks at work
ELF-MF JEM[35]	Extremely low-frequency magnetic fields

[a] Items differed slightly between AMIGO, EPIC-NL and the Nightingale questionnaires; more detailed information can be found in supplementary Table S1.
[b] More detailed information can be found in Schüz et al[14].
ELF-MF, extremely low-frequency magnetic fields; EMF, electromagnetic fields; HIT, Headache Impact Test; L_{den}, day-evening-night weighted average traffic noise level for 2011; LUR, land use regression; MEQ, Morningness-Eveningness Questionnaire; MOS, Medical Outcomes Study; NDVI, Normalized Difference Vegetation Index; NO_2, nitrogen dioxide; OP_{dtt}, oxidative potential; PAHs, polycyclic aromatic hydrocarbons; $PM_{2.5}$, particulate matter with diameter ≤2.5 µm; $PM_{2.5}$ absorbance, reflectance on $PM_{2.5}$ filters (marker of black carbon); PM_{10}, particulate matter with diameter ≤10 µm; RF-EMF, radiofrequency electromagnetic fields; SF, Short Form Health Survey.

In addition to the core LIFEWORK data, each subcohort had additional research aims and collected more extensive information on dedicated topics. In the Nightingale Study, this included shift work and more established hormone-related cancer risk factors, including reproductive factors[17].EPIC-NL was initiated to study the relationship between nutrition and cancer, and later the focus of the questionnaires broadened to study the etiology of other major chronic diseases and reproductive health[11]. The AMIGO questionnaire incorporated additional questions on self-reported health (e.g., respiratory health and somatization symptoms), indoor exposures, and risk perceptions and concerns about environmental exposures[12,21]. A LIFEWORK follow-up questionnaire campaign is expected to be complete by the end of 2017: it was administered for the EPIC-NL and AMIGO subcohorts from April until November 2015, and for the Nightingale subcohort starting in July 2017 and is expected to be completed by the end of 2017. The contents of the follow-up questionnaire are described in Supplementary Table S1. The intention is to continue to collect questionnaire data every 4-5 years, contingent on funding.

Exposure assessment

Detailed information was collected on occupational and residential histories, which allows for estimation of life course residential and occupational exposures. The various geospatial environmental exposure models and job-exposure matrices (JEMs) that have thus far been applied are listed in Table 1.

So far, several environmental exposures have been modeled for the baseline geocoded[36] home addresses. For example, radiofrequency (RF-)EMF from mobile phone base stations at the reported bedroom floor, where people spend the most time while at home, were estimated for each participant using the 3D radio wave propagation model, NISMap[23,24]. A measure of greenness in a 1000 m buffer around the home was assigned by calculating the mean of the Normalized Difference Vegetation Index (NDVI) for 2011. The NDVI is the difference between red and near infrared radiation, ranging from -1 to +1, where low values reflect low vegetation (e.g., water) and high values reflect high vegetation[27]. Ambient air pollution estimates [NO_2 (nitrogen dioxide), $PM_{2.5}$ and PM_{10} (particulate matter with an aerodynamic diameter ≤2.5 µm and ≤10 µm), $PM_{2.5}$ absorbance (a marker of black carbon), and oxidative potential of $PM_{2.5}$] were derived using the land-use regression (LUR) models developed for the European Study of Cohorts for Air Pollution (ESCAPE) for the years 2008-2011[25,26]. The models incorporate air pollution monitoring network, land use, and traffic count data. Exposure from road traffic noise was estimated with the Standard Model Instrumentation for Noise Assessments[28]; this model incorporates data on traffic intensities, speed, composition, type of road surface, building data, and ground type. 24-hr weighted average road traffic noise (L_{den}) levels, with penalties for the evening and night time periods, were estimated at the home addresses[28]. Estimates of PM element composition, including trace elements, and aspects of the built environment (e.g., number of nearby fast-food restaurants and sports facilities) will also be modeled[29,37].

Occupational exposures can be assigned based on the self-reported occupational histories, and linked to JEMs. JEMs are cross-tabulations of exposure scores and job titles that are used to assign quantitative or qualitative occupational exposure levels. As (expert-based) classification of job titles is a time consuming exercise, we decided to code a random subset allowing for more rapid implementation of the efficient case-cohort study design, as previously applied[38]. As of 2016, job histories prior to baseline (2011/2012) have been coded using the International Standard Classification of Occupations (ISCO-88) classification system for a random subset of 4961 participants. The job-specific codes have until now been linked to the ALOHA+ JEM, DOM-JEM, and shock-JEM to estimate lifetime exposure to numerous chemical, physical, and biological agents, as described in Table 1[30-34,39]. Within the aforementioned JEMs, limited distinction is made for the different nursing occupations or job tasks, as the ISCO-88 coding scheme has only a few specific codes for nursing jobs. Therefore, additional questions regarding specific occupational exposures were asked in the Nightingale questionnaire for nurses (e.g., contact with antineoplastic drugs, antibiotics).

As part of the COSMOS study, in addition to self-reported data, information on mobile phone use is obtained from mobile phone network operators for the participants who consented.[14] Operator data includes the monthly number and duration of calls, and number of SMS text messages. Operator data is continuously available from 3 months prior to the baseline questionnaire; so the collection period for operator and self-reported data overlaps. Operator data was only obtained for ~50% of participants who consented as we could not retrieve information for participants with e.g., subscriptions registered via an employer or prepaid service. We have so far obtained operator data for 12% (n=10 394) of the cohort for the baseline (2011-2012) period (from one operator). Data for a total of around 21% (n~18 500) is being retrieved for subsequent time periods and from additional operators.

Health outcome ascertainment

Multiple health outcomes were ascertained in the baseline questionnaires, incidence is monitored through regular linkage to registries, and changes in self-reported health are being assessed in follow-up questionnaires. Linkage to the registries was based on personal identifying information. Information on vital status was obtained from the Municipal Personal Records Database, and information on cause of death was obtained from the national mortality registry (Statistics Netherlands) in 2015–2017[40]. Data on cancer incidence was obtained from the Netherlands Cancer Registry; data is available from 1989 onwards and the most recent linkage was in 2017[41]. Most participants in AMIGO and EPIC-NL (33% of LIFEWORK) gave consent to retrieve health outcome information from their general practitioner. In AMIGO, information was obtained from the NIVEL Primary Care Database, which comprises electronic health record data from primary care physicians on consultations, drug prescriptions, and referrals from 2005 onwards[42,43]. Linkage to cause-of-death and cancer health registries will occur on a (bi)

annual basis, and linkage to other registries (e.g., the Hospital Discharge Registry) will be performed periodically, depending on research projects.

Health outcomes which are not easily captured by registries were assessed using validated questionnaires, such as headache (HIT-6), migraine (ID-migraine), sleep quality (MOS), general well-being (SF-12), and respiratory symptoms (based on the European Community Respiratory Health Survey)[16-19, 44,45]. As part of the individual study protocols, biological materials were gathered for a subset of the participants. Blood samples were collected for the majority of the EPIC-NL participants (in 1993-1997)[11]. In the Nightingale Study, toenail clippings were collected (n=23 439) on which DNA and other analyses can be performed[13].

Statistical analyses

We assessed the repeatability of several questionnaire items related to exposure assessment and a primary exposure of interest, mobile phone use, along with several characteristics we expected to remain stable (e.g., country of origin, height). The repeatability was assessed by percent observed agreement and reliability: weighted Cohen's kappa statistic for categorical data (equal κ_w for nominal data and squared κ_w for ordinal data) or intraclass correlation coefficient (ICC) for continuous data. We performed a stratified analysis on those who completed two questionnaires 1–4 months versus more than 5 months apart, which approximates the median time (151 days) between completing the two questionnaires; we excluded men and those who had completed the questionnaire within a period of less than one month (n=28). Minimal relative risks (RR) detectable with 80% power were estimated for the end of 2016 and 2021 by a logistic generalized additive model for different exposure prevalences (5-50%) for three diseases with varying age-standardized incidence and mortality rates. Information on the size, age and sex distribution of LIFEWORK was combined with outcome-specific incidence or mortality rates for the Netherlands[40]. Statistical analyses were performed using R, version 3.2.2 (R Foundation for Statistical Computing, Vienna, Austria). Geospatial modelling was performed using ArcGIS 10.3 (ESRI, Redlands, CA, USA).

FINDINGS TO DATE

Baseline characteristics

Baseline characteristics of the LIFEWORK cohort are shown in Table 2. A large proportion of the cohort is female (n=79 162, 90% female; n=9 304, 10% male), as only females were recruited for the Nightingale Study and one of the two EPIC-NL cohorts (Prospect)[11]. At LIFEWORK baseline (2011-2012), the median age was 51 years, and 43% were between 50-64 years of age. Three-quarters of participants were employed. Nearly half were overweight or obese, and more than half had ever smoked. 72% of

Table 2. Baseline characteristics of the LIFEWORK participants (n = 88 466, 2011–2012).

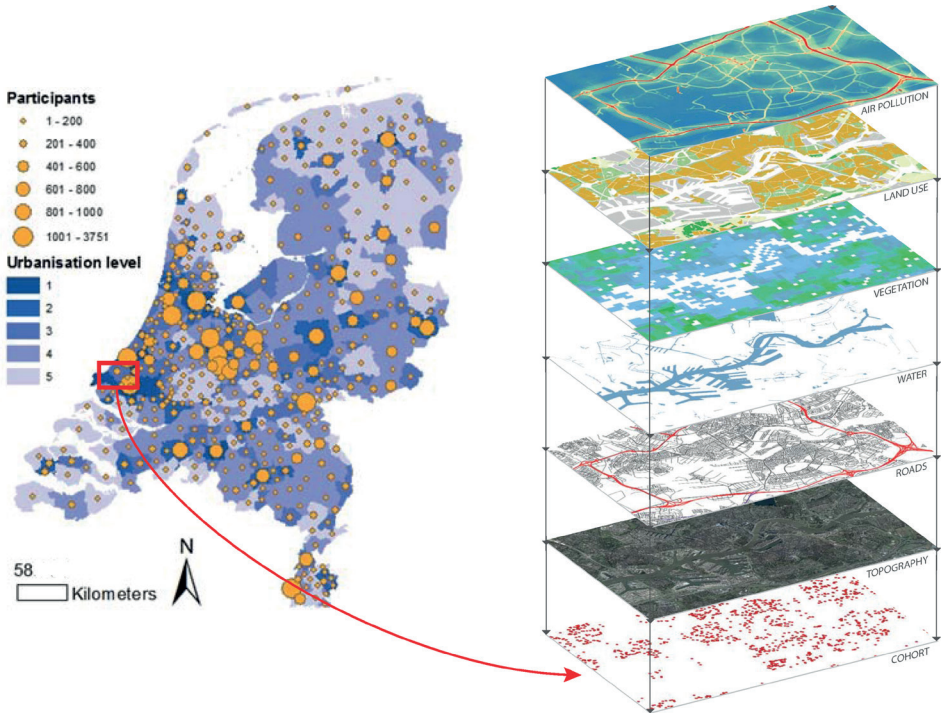
Characteristic	Women		Men		Total	
	N	%	N	%	N	%
N	79 162	90	9 304	10	88 466	100
Age (years), median, IQR	50, 41-58		54, 46-61		51, 42-59	
19–29	6 347	8	0	0	6 347	7
30–39	10 485	13	961	10	11 446	13
40–49	20 487	26	2 258	24	22 745	26
50–64	33 226	42	4 854	52	38 080	43
≥65	8 616	11	1 231	13	9 847	11
Missing	1	0	0	0	1	0
BMI (kg/m^2), median, IQR	24, 22-27		26, 24-28		25, 22-27	
Overweight (25–30 kg/m^2)	24 093	30	4 333	46	28 426	32
Obese (≥30 kg/m^2)	9 446	12	1 296	14	10 742	12
Missing	451	1	72	1	523	1
Marital status						
Married, registered partnership or living together	62 746	79	7 560	81	70 306	80
Divorced	4 362	6	440	5	4 802	5
Widow(er)	3 194	4	143	1	3 337	4
LAT relationship	2 140	3	164	2	2 304	3
Single	6 422	8	936	10	7 358	8
Missing	298	0	61	1	359	0
Country of birth						
The Netherlands	75 919	95.9	8 821	94.8	84 740	95.8
Indonesia[1]	535	0.7	88	0.9	623	0.7
Suriname	428	0.5	16	0.2	444	0.5
Germany	386	0.5	47	0.5	433	0.5
Belgium	251	0.3	26	0.3	277	0.3
Former Netherlands Antilles[2]	105	0.1	54	0.6	159	0.2
Other	1 426	1.8	252	2.7	1 678	1.9
Missing	112	0.2	0	0	112	0.1
Level of urbanisation[3]						
Very high	11 394	14	1 562	17	12 956	15
High	17 896	23	2 041	22	19 937	22
Moderate	15 443	19	1 689	18	17 132	19
Low	17 119	22	1 947	21	19 066	22
Very low	16 821	21	2 018	22	18 839	21
Missing	489	1	47	0	536	1
Monthly income estimate[4]						
Low	1900	2	344	4	2244	3
Medium	60970	77	7228	78	68198	77
High	13038	17	1353	14	14391	16
Missing	3254	4	379	4	3633	4
Highest level of education attained[5]						
Low	9 194	12	2 789	30	11 983	14
Intermediate	37 272	47	2 743	30	40 015	45
High	32 506	41	3 765	40	36 271	41
Missing	190	0	7	0	197	0

Characteristic	Women		Men		Total	
	N	%	N	%	N	%
Employment status						
Self-employed	3 394	4	1 210	13	4 604	5
Employed	55 309	70	5 348	58	60 657	69
Retired	7 124	9	1 739	19	8 863	10
Unemployed	643	1	215	2	858	1
Sick leave/disability	2 309	3	407	4	2 716	3
Other: stay-at-home parent/voluntary work/ student, etc.	8 463	11	249	3	8 712	10
Missing	1 920	2	136	1	2 056	2
Smoking status						
Never	37 182	47	3 856	41	41 038	46
Former	31 318	40	3 995	43	35 313	40
Current	9 700	12	1 357	15	11 057	13
Missing	962	1	96	1	1058	1
Alcohol consumption[6]						
Never	5 344	7	199	2	5 543	6
Current	70 842	89	8 611	93	79 453	90
Ever	2 774	4	478	5	3 252	4
Missing	202	0	16	0	218	0
Mobile phone use						
Ever	56 621	72	7 034	75	63 655	72
Never	20 817	26	2 032	22	22 849	26
Missing	1 724	2	238	3	1 962	2
Mobile phone use in past 3 months[7]						
<5 min/week	11 648	20	1 149	16	12 797	20
5–29 min/week	23 539	42	2 593	37	26 132	41
30–59 min/week	9 260	16	1 378	20	10 638	17
1–3 h/week	7 155	13	1 078	15	8 233	13
4–6 h/week	1 849	3	474	7	2 323	4
≥6 h/week	1 122	2	362	5	1 484	2
Missing	2 048	4	0	0	2 048	3
Years of mobile phone use[8]						
≤4 years	3 330	6	212	3	3 542	6
5-9 years	25 282	45	2 927	42	28 209	44
10-14 years	7 143	13	2 225	32	9 368	15
≥15 years	12 564	22	876	12	13 440	21
Missing	8 302	14	794	11	9 096	14

[1] Indonesia including former Dutch East Indies.
[2] Aruba, Bonaire, Curacao, St. Martin, St. Eustatius.
[3] Average number of addresses/km^2 on postal code level within a radius of 1 km of the home address at LIFEWORK baseline; categorized into five levels ranging from very high=on average >2500 addresses/km^2; high=on average 1500–2500 addresses/km^2; moderate=on average 1000–1500 addresses km^2; low=on average 500–1000 addresses/km^2; and very low=on average <500 addresses/km^2.[40]
[4] Household income was estimated based on participants' baseline postal code. Each postal code was linked to income data from Statistics Netherlands for December 2008; the cutoff values for low and high income were respectively the 40th and 80th percentile in income.[40]
[5] Low: primary school, lower vocational training or lower secondary education; intermediate: intermediate vocational education or intermediate/higher secondary education; high: higher vocational education or university degree.
[6] Current: more than 1 glass per week in the past 12 months; ever: less than 1 glass per week in the past 12 months.
[7] Calculated among the participants that reported using a mobile phone in the 3 months prior to baseline (n=57 644).
[8] Calculated among the participants that reported having ever used a mobile phone (n=63 655) for the year 2012.
BMI, body mass index; IQR, interquartile range; LAT, living-apart-together.

participants reported ever having used a mobile phone. Over a third (36%) reported to call >30 min/ week in the 3 months prior to baseline. Twenty-two percent (or 36% of the participants who had ever used a mobile phone) had used a mobile phone for more than 10 years by 2012. Participants resided in all areas of the Netherlands (see Figure 1), although some clustering can be identified (such as in the central province of Utrecht, where EPIC-NL recruitment was concentrated), and slightly more participants lived in rural areas than urban areas.

Figure 1. Distribution of LIFEWORK participants (circles) and urbanisation level per municipality across the Netherlands.



Five urbanisation levels ranging from 1 (low, <500 addresses/km^2) to 5 (very high, ≥2500 addresses/km^2). Overlay of environmental layers in the Rotterdam area, respectively from bottom to top: location of participants, topography, roads, water, greenness (Normalized Difference Vegetation Index), land use, and air pollution.

Environmental exposure distributions

Exposure distributions for a selected set of residential environmental exposures are shown in Table 3. The median exposure level was 22.9 µg/m^3 for NO_2, 16.6 µg/m^3 for $PM_{2.5}$, 53.1 dB for noise, and 0.003 mWm^2 for RF-EMF from mobile phone base stations. Contrast (range/mean) for $PM_{2.5}$ and NO_2 was 39% and 339%, respectively. Spearman correlation coefficients between the environmental exposures

ranged from 0.14 to 0.86 (Supplementary Figure S1). The highest correlations were found between several air pollutants (0.86 for $PM_{2.5}$ absorbance and PM_{10}), and negative correlations ranging from -0.20 to -0.54 were observed between greenness (NDVI) and all other environmental exposures.

Table 3. Exposure distributions for environmental exposures at the baseline home addresses based on geospatial modeling (LIFEWORK, n = 88 466, 2011–2012).

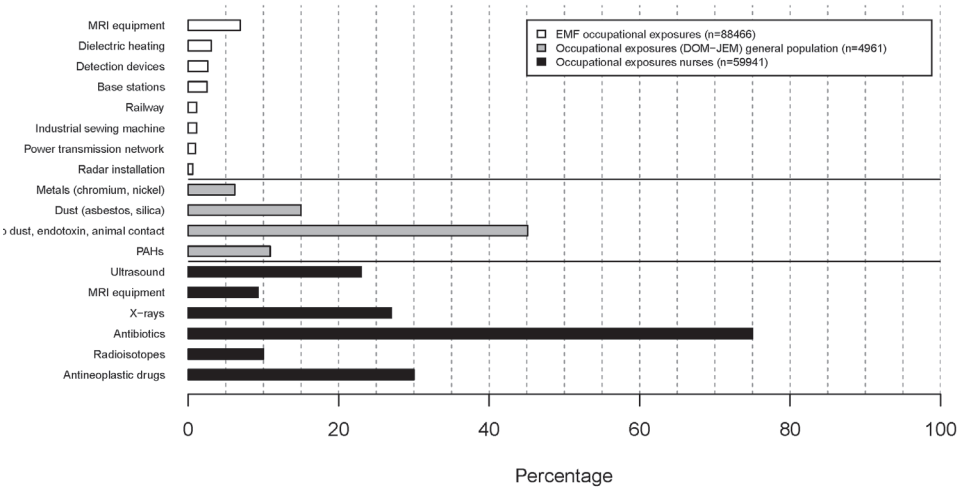
	10th P	**25th P**	**50th P**	**75th P**	**90th P**
Outdoor air pollution[1]					
NO_2 (µg/m^3)	16.05	19.07	22.93	27.35	32.01
PM_{10} (µg/m^3)	23.80	23.97	24.43	25.15	26.18
$PM_{2.5}$ (µg/m^3)	15.58	16.17	16.57	17.04	17.32
$PM_{2.5}$ absorbance (10^{-5} m^{-1})	0.99	1.12	1.23	1.37	1.53
OP_{dtt} (nmol DTT/min/m^3)	0.88	1.04	1.19	1.31	1.41
Mobile phone base station RF-EMF[2]					
RF-EMF (mWm^2)	0.000	0.000	0.003	0.017	0.063
Greenness in 1000 m buffer[3]					
NDVI (scale from -1 to + 1)	0.41	0.47	0.52	0.58	0.64
Traffic noise exposure[4]					
L_{den} (dB)	47.00	49.80	53.10	57.10	61.50

[1] LUR ESCAPE model based on data from 2008–2011; [2] NISMap estimates for 2011 or 2012, depending on the year the baseline questionnaire was filled in; [3] NDVI in 2011; [4] STAMINA in 2011.
L_{den}, day-evening-night weighted average traffic noise level; LUR, land-use regression; NDVI, Normalized Difference Vegetation Index; NO_2, nitrogen dioxide; OP_{dtt}, oxidative potential measured by dithiothreitol; P, percentile, $PM_{2.5}$, particulate matter with diameter ≤2.5µm; $PM_{2.5}$ absorbance, reflectance on $PM_{2.5}$ filters (marker of black carbon); PM_{10}, particulate matter with diameter ≤10µm; RF-EMF, radiofrequency electromagnetic fields.

Occupational exposures

The LIFEWORK baseline prevalence of selected occupational exposures is shown in Figure 2 for sources of occupational-related EMF exposures (in the entire cohort, n=88 466), occupational exposures based on the DOM-JEM for a subset of the AMIGO general population cohort (n=4 961), and for nurse-specific occupational exposures (Nightingale, n=59 941). The most common sources of recent occupational-EMF exposure was working in the vicinity of magnetic resonance imaging equipment (7%) and using dielectric heating (3%). In the subset for whom occupational histories were linked to DOM-JEM, lifetime exposures were highest (25%) for organic dust (a mixture of dust from organic substances like wood, flour, etc.) and endotoxin (16%). The majority of the Nightingale Study participants reported to have ever worked with antibiotics (75%), and around a quarter (23-27%) with antineoplastic drugs, routine X-rays, or ultrasound equipment. A large proportion of LIFEWORK participants, 66%, reported to have ever conducted shift work; 80% of Nightingale and 32% of the two general population cohorts (AMIGO and EPIC-NL).

Figure 2. Frequency of selected lifetime occupational exposures based on self-reports or estimated from job-exposure matrices.



These are stratified by exposures to occupational sources of high electromagnetic fields at baseline (n=88 466, i.e., all LIFEWORK participants); lifetime exposure to occupational chemicals (DOM-JEM[32, 33]) in the subset with coded job-titles (n=4 961 in AMIGO); and lifetime exposure for at least 6 months to occupational agents in the subcohort of nurses (n=59 941, the Nightingale Study). EMF, electromagnetic fields; MRI, magnetic resonance imaging; PAHs, polycyclic aromatic hydrocarbons.

Repeatability of questionnaire items

For the 265 participants who completed the LIFEWORK questionnaire twice, the median interval between completing the two questionnaires was 151 days (interquartile range (IQR): 122-186). These participants had a mean age of 51 years (IQR 30-66) and 98% were female. The repeatability of several exposures and covariates is presented in Table 4. Repeatability of country of origin and height was excellent ($\kappa_w \geq 0.97$). Questions on recent mobile phone use (duration) had a moderately good repeatability ($\kappa_w = 0.73$). For historical use (duration), the weighted kappa was poor to moderate at 0.04, 0.21, 0.70, 0.69, and 0.55 respectively for 1990, 1995, 2000, 2005, and 2010 (Supplementary Table S2). In stratified analyses, no clear differences were observed for participants who filled in the questionnaires between 1-4 months versus more than 5 months apart.

Table 4. Test-retest repeatability[1] for selected key items in the LIFEWORK baseline questionnaire.

Variable	Overall			1-4 months			≥ 5 months		
	N	Percent Agreement	Reliability (κ_w or ICC)	N	Percent Agreement	Reliability (κ_w or ICC)	N	Percent Agreement	Reliability (κ_w or ICC)
General									
Country of origin	237	100	1	83	100	1	154	100	1
Height (cm)[2]	237	61.6	0.97	83	67.5	0.99	154	58.4	0.97
Since when at current address (year)[3]	236	78.6	0.98	82	84.5	0.99	154	74.7	0.97
Floor of bedroom[3]	231	87.4	0.90	81	87.7	0.72	150	87.3	0.92
Bedroom window glazing[3]	233	95.3	0.84	81	98.8	0.96	152	93.4	0.75
Mobile phone use[4]									
Past 3 months: mobile phone use (duration, categorical)	107	57.9	0.73	32	56.2	0.62	75	58.7	0.76
Mobile phone use in 2005 (duration, categorical)	107	59.8	0.69	32	65.6	0.65	75	57.3	0.70
Mobile phone use in 2000 (duration, categorical)	107	57.0	0.70	32	50.0	0.52	75	60.0	0.75
Laterality (held on left, right, equal)	138	78.3	0.69	38	84.2	0.76	100	76.0	0.66

[1] Excluding men (n=1) and participants filling in the questionnaire within one month (n=27).
[2] Height for the EPIC-NL subcohort is based on baseline in EPIC-NL which was in 1993–1997.
[3] Excluding the participant that moved in the period between completing the LIFEWORK baseline questionnaire twice (n=1).
[4] Only calculated among participants that reported to use a mobile phone at baseline (percentage agreement = 82.1, n=235).

Power calculations

Minimal detectable for the different exposure scenarios (5-50% exposure prevalence) and years (2021 and 2016) RRs ranged from 1.12 to 1.28 for stroke (age-standardized mortality: 0.700/1000), 1.36 to 2.36 for Parkinson's disease (age-standardized incidence: 0.074/1000), and 1.64 to 3.60 for brain cancer (age-standardized mortality: 0.006/1000) (Supplementary Figure S2). Self-reported doctor diagnosis of conditions and disease prevalence at baseline is shown in Supplementary Table S3, and was the highest for high blood pressure, cholesterol, migraine, and asthma (18.8-6.5%).

STRENGTHS AND LIMITATIONS

Through an efficient approach of collecting data in three different cohort studies, we established the LIFEWORK cohort, the largest contemporary Dutch prospective cohort study on occupational and environmental risks with a nationwide distribution in the Netherlands. The detail and wealth of the data collected on a broad array of occupational and environmental exposures is illustrated in this paper. The data enables not only assessments of health effects of single exposures, but also of more complex epidemiological analyses on multiple simultaneous and time-varying exposures, such as the combined effects of electromagnetic fields, green space, air pollution, traffic noise, and work-related exposures. In addition to extensive exposure data, information on a large set of potential confounders and effect modifiers, and self-reported and registry-recorded acute and chronic health outcomes is available.

The design of LIFEWORK was efficient but influenced the composition of the cohort; for example, the majority of the cohort is female (90%), over the age of 50 years (54%), and nursing is the most common occupation. There will be greater statistical power to study associations with female-specific health outcomes, such as breast cancer and diseases which are more prevalent in women, such as certain autoimmune diseases and mental disorders (e.g., major depression)[46]. There should be sufficient statistical power to study effect modification by sex for many exposure–outcome associations, but likely not for some rarer exposures or outcomes. Furthermore, exposures associated with the nursing profession will be more prevalent in the study population compared to the general working population. The participants are higher educated than the general population because nurses, who account for more than 60% of the study population, have at least an intermediate level of education[40]. Other consequences of the recruitment strategies is clustering of participants in the center of the country (in and around Utrecht), and that the proportion of participants employed at baseline is high (74% versus 71% for the adult population of the Netherlands)[40], which implies high statistical power for investigating occupational risks, especially for exposures related to nursing.

That LIFEWORK is not fully representative of the Dutch adult population with respect to sex, age, and occupation, does not hamper the ability to detect and estimate exposure-outcome associations and for valid inferences, assuming adequate control of confounding variables[47]. We have collected data on a diverse set of potential confounding variables and endeavored to minimize their measurement error. A limitation, common to many observational prospective cohorts, is selection bias, particularly due to attrition. This will be minimized by using health registries to assess health outcomes, and may, in case of usage of questionnaire data, also be accounted for using modelling approaches to account for selection effects and time-varying confounding[48-50].Due to the prospective assessment of health outcomes, possible misclassification of exposures based on retrospective assessment of lifetime occupational, residential and mobile phone histories is expected to be largely non-differential. Healthy-worker survivor effect is a concern in occupational cohorts; however, for the Nightingale Study, participants were recruited from the healthcare professional registration system and working in the field was not a prerequisite for inclusion. Both AMIGO and EPIC-NL recruited from the (from the (working age) general population), for which healthy-worker effect is less of a problem.

Two of the three subcohorts were invited to complete a web-based questionnaire, and one subcohort (EPIC-NL, 16%) with an older age distribution, was invited to complete a paper questionnaire which was the customary approach in that study. An advantage of the web-based questionnaire was that more participants could be invited due to the lower cost of printing and data-entry[51]. This might have led to an overrepresentation of more computer proficient individuals (generally more educated and younger); however, we expect this selection to have been minimal as internet access in the Netherlands is very high (>97%)[52], and no difference in response rates was observed in a recent study of older patients in the Netherlands which offered web-based questionnaires (with paper optional) and paper-based questionnaire[53]. Minor differences in the completion rates and data quality can be expected for the web-based versus paper questionnaires, for example because error messages appeared in the web-based versions when sections were incomplete or highly unlikely values were entered. Further, web-based questionnaires may be less prone to social desirability bias[54].

LIFEWORK is the second largest contributor to the COSMOS study, comprising nearly 90 000 of the 290 000 participants enrolled to date.[14] Mobile phone use was somewhat lower in LIFEWORK compared to the study populations from the other countries in COSMOS, especially the proportion who frequently called with a mobile phone (>30 min/week). Explanations for this are likely the older age distribution of LIFEWORK compared to the other countries' cohorts, and that the other COSMOS countries oversampled, based on operator records, heavy mobile phone users, and excluded those not using mobile phones (which constituted 26% of LIFEWORK participants). A strength of COSMOS is that in addition to gathering questionnaire data on mobile phone use, objective data is collected from operators, allowing for application of measurement error models (e.g., regression calibration).

Another strength is the prospective cohort design, which reduces possible recall bias and selection bias. This was a possible limitation of previous case-control studies on mobile phone use and health[55]. Compared to the large INTERPHONE study on mobile phone use and brain tumors, participants of LIFEWORK (and cohorts comprising the COSMOS consortium) have a longer time since first use and a higher proportion are heavy users[56]. 26% of LIFEWORK participants reported using a mobile phone >30 min/week at baseline and the average time since first usage was 11 years (IQR 8-14 years), which is higher than the highest category in the Interphone study (≥10 years)[56]. This will be informative to address current scientific uncertainties regarding the possible health effects long-term and heavy mobile phone usage, although cumulative RF-EMF exposures will in part be compensated by the fact that modern phones produce less RF-EMF due to adaptive power control and network evolution (e.g., 2G to 3G)[2]. The detailed information on the use of mobile communication devices also allows for research on borderless working and (blue) light-at-night exposure, exposures of emerging concern.

Because 265 participants completed the baseline LIFEWORK questionnaire twice, we assessed the repeatability of certain questionnaire items, such as questions related to mobile phone use, which is one of the key research areas of interest. The repeatability of questions on recent mobile phone use was substantial. The low kappa for 1990 and 1995 for mobile phone usage (duration) was probably due to the low overall usage in these years; however, the percentage agreement for these periods was high (90% and 79%). Stratified analysis showed that the time between the two questionnaires (more or less than 5 months) did not result in consistently different repeatability coefficients. Variables answered with substantial repeatability, such as mobile phone usage, will increase statistical power, while variables with moderate agreement, such as bedroom floor (important for modelling some environmental exposures), will have greater measurement error and consequently reduced statistical power.

A large amount of data has been collected for nearly 90 000 participants in LIFEWORK. The large sample size, widespread geographic coverage of participants across the Netherlands, and the large contrast in levels of occupational and environmental exposures will allow us to study both common diseases and relatively rare diseases, such as Parkinson disease, as shown by the power analysis. This wealth of data creates opportunities but will also lead to analytical challenges, such as disentangling the effects of duration and rate of exposure histories, and the independent effects of multiple, potentially correlated exposures[57-59]. Pooling resources to create the LIFEWORK cohort enabled us to design a larger cohort with a broad array of exposures and outcomes; however, there are challenges to such a federated structure. Reaching consensus on core questionnaire topics can be resource intensive, although we do benefit from diverse expertise. Furthermore, coordinating follow-up of the cohort is a challenge, and ultimately there will likely be some time gaps between some of the questionnaire campaigns across the subcohorts and potentially unequal coverage of registry linkage.

Future plans

We are now in the phase of actively and passively observing participants; data for the first follow-up questionnaire has been collected for two of the three subcohorts (in 2015), and is expected to be completed for the third subcohort by the end of 2017, and linkage to cause-of-death and cancer registries has been performed for all subcohorts as of 2017. We plan to administer questionnaires every 4-5 years, and aim to continue this for 20+ years, depending on future funding. As we have the email addresses of ~80% of the cohort, it would be possible to contact a subset of the cohort in a cost-efficient manner. Using web-based invitations and questionnaires would eliminate printing and mailing costs, and reduce data management costs; nonetheless, costs associated with modifying the online questionnaire and providing support to participants can still be substantial. We could target specific groups (e.g., higher exposed participants), or send dedicated in-depth questions based on their answers in earlier questionnaires. We are also piloting the use of smartphone applications to passively monitor exposures—presently mobile phone usage, with plans to extend this to ecological momentary assessments of time-activity, and food frequency surveys—and are exploring using wearable passive samplers to assess exposures and physical activity in a subset of participants[60,61].

Collaboration

The LIFEWORK prospective cohort will continue to collect a rich set of data on multiple exposure domains and health outcomes. For more information refer to the website: lifeworkstudy.nl. Researchers interested in collaboration are invited to propose occupational and environmental research based on the data available within LIFEWORK or to submit a request for additional data collection. Requests can be submitted to R. Vermeulen [r.c.h.vermeulen@uu.nl] and will be reviewed by the LIFEWORK scientific board. LIFEWORK has an overarching governance board, and each subcohort has an independent and partially overlapping governance board. It is also possible to seek independent collaboration with one of the subcohorts (refer to http://www.amigoproject.nl/contact/; https://www.epicnl.eu/Home/EPICNL; and http://www.nightingalestudie.nl/Pages/for-researchers/ for contact information).

Acknowledgements

The authors wish to thank Eef van Otterloo from Utrecht University, Susan van Hemert from the University Medical Center Utrecht and Jan van der Laan from the Dutch National Institute for Public Health and the Environment (RIVM) for helping with the data collection, online applications, technical support, and data management tasks. They also are grateful to Joris IJzermans (NIVEL) and the general practitioners for their contribution to the establishment of AMIGO. The authors also thank Jeroen van Leuken, Danny Houthuijs and Wim Swart from the Dutch National Institute for Public Health and the Environment (RIVM) and Maciej Strak from Utrecht University for their help in providing

and modelling the greenness, noise and air pollution estimates. They also thank Lützen Portengen from Utrecht University for the power analysis simulation, Anke Huss and Susan Peters from Utrecht University for their contributions to the development of several job exposure matrices, and thank Mattijs Numans of the Utrecht Health Project (Leidsche Rijn Gezondheids Project) for his contribution to the study in the initial phase.

Funding

This research was funded by the Netherlands Organisation for Health Research and Development (ZonMw) within the Electromagnetic Fields and Health Research program (grant number 85200001 and 85500003), the National Institute for Public Health and the Environmental (RIVM) and the Netherlands Cancer Institute. The EPIC-NL study was funded by “Europe against Cancer” Programme of the European Commission (DG SANCO); the Dutch Ministry of Health, Welfare and Sports (VWS); the Netherlands Organisation for Health Research and Development (ZonMw); and the World Cancer Research Fund (WCRF). ESCAPE received funding from the European Community’s Seventh Framework Program (FP7/2007-2011) under grant agreement number 211250.

SUPPLEMENTARY TABLES AND FIGURES CHAPTER 2

Table S1. Overview of data collected for the LIFEWORK study, including differences in data collection per subcohort

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			Comments
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	
QUESTIONNAIRE								
General								
General (e.g., sex, height, weight, handedness)	√	√		Y	√	√	√	Only characteristics that might be expected to change (e.g., weight) were asked in follow-up.
Demographic information (country of birth, education, marital status)	√	√		H	√	√	√	Minor differences in country of birth, marital status and education questions.
Lifestyle (alcohol use, smoking)	√	√		H	√	√	√	Minor differences in questions; harmonized categories have been created.
Physical activity	√	√		H	√ (F)	√	√	
Basic diet				N	√ (F)			
Extended food frequency questionnaire				N		√ (F)		
Waist-hip circumference (self-reported)		√		Y	√ (F)	√ (B/F)	√ (F)	Only collected for EPIC-NL at baseline (2011); for all 3 subcohorts at follow-up.
Residence								
Bedroom floor	√	√		Y	√	√	√	

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	Comments
Bedroom, extended (window opening habits, if streetside)	√	√		N	√	√	√ (B)	
Domestic/indoor environment (incl. cooking, ventilation)	√	√		N	√	√		
Historical addresses (those resided at for at least 12 months)	√			Y	√ (B)	√ (B)	√ (B)	Address changes after baseline are obtained from municipal records.
Green space (proximity to and use of)		√		N	√ (F)	√ (F)		
Noise exposure, sensitivity[1]		√		N	√ (F)	√ (F)		
Light at night exposure in bedroom		√		N	√ (F)	√ (F)	√	
Use of electronics (light at night) in hour before sleep		√		Y	√ (F)	√ (F)	√ (F)	
Employment								
Current (job, tasks, hours, physical exertion)	√	√		H	√	√	√	Asked in follow-up only if participant changed occupation since baseline.
Historical occupations (at least 6 months)	√	√		H	√	√	√	The complete history of jobs only asked at baseline; participants asked to update at follow-up if job changed since baseline.
Shift and night work[d] (ever shift work; if permanent or rotating)	√	√		H	√	√	√	Shift and night work was more comprehensively assessed in the Nightingale questionnaire

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			Comments
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	
Extended information on shift work				N			√	Start and stop time of shift, rotating or permanent, and speed and direction of a rotating system, years on a particular non-day shift schedule and cumulative exposure to the shift system over the subject's working life, and shift intensity (i.e. time off between successive work days on the shift schedule).
Work satisfaction and job stress[2]				N	√			
Occupational EMF exposures								
Induction/electric welding, induction/ electric/gas oven, induction/electric hob, MRI, detection gate, radar installations (last 3 months (EPIC/AMIGO), last 6 months (Nightingale))	√	√		H	√	√	√	Only questions relevant for nurses asked in Nightingale.
Mobile phone use history[3]								
First use (year)	√			Y	√	√	√	
Frequency of voice calls in 5 year period (categorical)	√	√		Y	√	√	√	
Duration of voice calls in 5 year period (categorical)	√	√		Y	√	√	√	

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	Comments
Hands-free usage in 5 year period (categorical)	√	√		Y	√	√	√	
Mobile phone use in past 3 months[3]								
Frequency and duration of voice calls (categorical)	√	√		Y	√	√	√	
Hands-free usage (categorical)	√	√		Y	√	√	√	
Mobile internet usage (categorical)	√	√		Y	√	√	√	
Smartphone usage (categorical)		√		Y	√ (F)	√ (F)	√ (F)	
Text messaging (categorical)	√	√		Y	√	√	√	
Internet usage (categorical; during the week and weekend)	√	√		N	√	√		
Tablet, computer usage (categorical)		√		N	√ (F)	√ (F)		
Cordless phones[3]								
First use at home (year)	√			Y	√	√	√	
First use at work (year)	√			Y	√	√	√	
Frequency of voice calls in past 3 months (categorical)	√	√		Y	√	√	√	
Duration of voice calls in past 3 months (categorical)	√	√		Y	√	√	√	
Environmental exposures								
Concerns about environmental exposures and health				N	√			
Perceived environmental exposures and health risks				N	√			
Health and well-being								
General well-being (SF-12)[4]	√	√		Y	√	√	√	
Headaches (HIT-6)[5], migraine (ID-Migraine)[6]	√	√		Y	√	√	√	Migraine questions more extensive in EPIC-NL follow-up.
Hearing and tinnitus	√	√		Y	√	√	√	

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	Comments
Sleep habits and problems (MOS Sleep Scale)[7]	√	√		Y	√	√	√	
Chronotype MCTQShift[8]/(MEQ)[9]				N		√ (B)	√	MEQ was assessed in EPIC-NL and Nightingale baseline. MCTQShift was assessed in Nightingale follow-up.
Recollection/memory				N	√			
Respiratory symptoms: (European Community Respiratory Health Survey)[10]	√	√		N	√	√		
Somatization symptoms (Four-Dimensional Symptom Questionnaire)[11]				N	√			
Stress (Sheldon Cohen Scale)[12]		√		N	√ (F)	√ (F)		
Historical and prevalent chronic diseases/ conditions (self-reported doctor diagnosis and age at diagnosis)[e]	√	√		Y	√	√	√	See Table S3 for the prevalence rates for diseases/ conditions that were assessed in all subcohorts, and footnote (e) for a list of conditions additionally assessed in specific subcohorts.
Medications (selected[f])	√	√		H	√	√	√	See footnote (f).
Parkinson's disease screening[13]	√	√		N	√	√		
Family history of disease[g]	√	√		H	√	√	√ (B)	See footnote (g).
Reproductive health (women only)								
Menstruation	√	√		N	√ (F)	√	√	

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			Comments
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	
Pregnancies	√	√		H	√	√	√	
Hormone use (incl. oral contraceptives, hormone replacement therapy)	√	√		N	√ (F)	√	√	
Fertility (incl. infertility treatments, time to pregnancy)	√	√		N	√ (F)	√	√	Time to pregnancy added at follow-up. Some questions (e.g., time to pregnancy) were not asked in EPIC-NL because this population was older (much time had elapsed pregnancy).
Birth outcomes	√	√		H	√ (F)	√	√	
OTHER DATA SOURCES								
Mobile phone traffic data from operators								Data is available for the subset that consented and were matched (~20%).
Frequency and duration of voice calls	√	√	√ (monthly)	Y	√	√	√	
Frequency of text messages	√	√	√ (monthly)	Y	√	√	√	
Data use volume (kB)	√	√	√ (monthly)	Y	√	√	√	
Routine health data (long-term follow-up)								
Vital status (Municipal Personal Records Database; in Dutch: Basisregistratie personen) and cause of death registry (Statistics Netherlands)			√ ((bi) annually)	Y	√	√	√	WHO coding guidelines: Only one underlying cause of death is recorded; secondary causes are also recorded.

Data	Baseline[a] (2011/2012)	Follow-up[a] (2015–2017)	Linkage (frequency)	Same data in all 3[b] (Y/N/H)	Differences across 3 subcohorts[c]			
					AMIGO[4] (n=14829)	EPIC-NL (n=13961)	Nightingale (n=59941)	Comments
Netherlands Cancer Registry (including topography, morphological classification, date of diagnosis)			√ ((bi) annually)	Y	√	√	√	Classification: International classification of diseases for oncology (ICD-O). Data are available from 1989.
Hospital Discharge Registry (In Dutch: LBZ; previously LMR)			√ (periodically, depending on research projects)	N		√		Hospital admissions, outpatient consultation and emergency department visits (ICD-9 coding); data from 1995 onwards; coverage >80% of Dutch population.
General Practitioner database (NIVEL Primary Care Database)			√ (periodically, depending on research projects)	N	√			Diagnoses coded following International Classification of Primary Care (ICPC). Available for AMIGO only. EPIC-NL participants' general practitioners can be approached with questionnaires.

B, baseline; F, follow-up questionnaire; HIT, Headache Impact Test; MCTQShift, Munich ChronoType Questionnaire for Shift-Workers; MEQ, Morningness-Eveningness Questionnaire; MOS, Medical Outcomes Study; MRI, magnetic resonance imaging; SF, Short Form Health Survey.

[a] Baseline questionnaires were completed in 2011–2012. The first (comprehensive LIFEWORK) follow-up questionnaire was conducted in 2015–2017 to obtain repeated measures on the exposures and health outcomes; some additional topics were assessed at follow-up, as indicated. If a box is checked, data was collected in at least two subcohorts.

[b] Y/N/H indicates whether the data collected (in questionnaires or via registry linkages) was collected in exactly the same way (Y) across all three subcohorts; was collected but using different questions or collection methods and was (or can later be) harmonized following a standardized protocol (H); or was not collected across all three subcohorts of LIFEWORK (N).

[c] If checked data is collected at both baseline and follow-up in this cohort. B/F indicates whether the data was collected at only baseline (B) or only follow-up(F)

[d] AMIGO: two extra questionnaires where sent out in 2013 and 2014 to a subset of the population with respect to symptoms, environmental risk and perception. Refer to Martens et al.[14] for a description.

[e] Diseases/conditions assessed at baseline: asthma, Chronic Obstructive Pulmonary Disease (COPD), cancer (differences between cohorts), angina pectoris, heart attack, other heart disease, stroke, high blood pressure,, diabetes, high cholesterol, chronic liver disease, thyroid disease, chronic inflammatory bowel disease (Crohn's disease/ colitis ulcerosa, only AMIGO/Nightingale), stomach ulcer (only AMIGO/Nightingale), chronic fatigue syndrome, depression, Attention Deficit Hyperactivity Disorder (ADHD, only AMIGO/Nightingale), hearing problems/deafness, cataract, glaucoma, migraine, concussion, epilepsy, Parkinson's disease, Amyotrophic Lateral Sclerosis (ALS), Multiple Sclerosis (MS), Dementia/Alzheimer's disease, auto-immune disease (only AMIGO/Nightingale). More extensive questions on cancer [breast, lung, bowel, ovarian, uterine, cervical, endometrial, Hodgkin's lymphoma, non-Hodgkin's lymphoma, leukemia, skin cancer (melanoma, basal/squamous cell), pancreatic] were assessed in the Nightingale baseline questionnaire.

[f] Medication: aspirin, painkillers, medication for high blood pressure, medication for high cholesterol, asthma medication, sleep medication (Nightingale only), Parkinson disease medication, COPD medication (AMIGO/EPIC-NL only), diabetes medication (insulin/tablets), antidepressants, immunomodulators (Nightingale only). More extensive set of questions on hormone medications asked in Nightingale questionnaire.

[g] Family history asked for father/mother (AMIGO/EPIC-NL/Nightingale baseline and AMIGO/EPIC-NL follow-up), brothers/sisters (EPIC-NL/Nightingale baseline, AMIGO/ EPIC-NL follow-up), sons/daughters (Nightingale baseline, AMIGO follow-up): cancer, diabetes, Parkinson's disease, Alzheimer's disease (not baseline EPIC-NL/AMIGO), stroke, heart attack, asthma, hay fever. More extensive questions on cancer in Nightingale (aunts' and grandparents' incidence of breast cancer, ovarian cancer, cervical cancer).

Table S2. Test-retest repeatability (maximum n=237) for mobile and cordless phone items in the LIFEWORK baseline questionnaire[a]

Variable	Overall			1–4 months[b]			≥ 5 months[b]		
	N	Percent Agreement	Reliability (κ_w or ICC)	N	Percent Agreement	Reliability (κ_w or ICC)	N	Percent Agreement	Reliability (κ_w or ICC)
Past mobile phone use									
Mobile phone use 2010 (frequency, categorical)	109	62.4	0.53	31	71.0	0.64	78	59.0	0.51
Mobile phone use 2005 (frequency, categorical)	109	61.5	0.67	31	71.0	0.59	78	57.7	0.69
Mobile phone use 2000 (frequency, categorical)	109	62.4	0.66	31	58.1	0.51	78	64.1	0.70
Mobile phone use 1995 (frequency, categorical)	109	78.0	0.33	31	74.2	0.52	78	79.5	0.27
Mobile phone use 1990 (frequency, categorical)	109	90.8	0.50	31	87.1	0.61	78	92.3	0.47
Mobile phone use 2010 (duration, categorical)	107	51.7	0.55	32	56.2	0.80	75	49.3	0.47
Mobile phone use 2005 (duration, categorical)	107	59.8	0.69	32	65.6	0.65	75	57.3	0.70
Mobile phone use 2000 (duration, categorical)	107	57.0	0.70	32	50.0	0.52	75	60.0	0.75
Mobile phone use 1995 (duration, categorical)	107	79.4	0.21	32	75.0	0.34	75	81.3	0.18
Mobile phone use 1990 (duration, categorical)	107	90.4	0.04	32	90.0	0.40	75	90.5	-0.04
Hands-free device use 2010 (categorical)	100	71.0	0.41	28	82.1	0.45	72	66.7	0.40
Hands-free device use 2005 (categorical)	100	80.0	0.45	28	89.3	0.46	72	76.4	0.45
Hands-free device use 2000 (categorical)	100	89.0	0.42	28	89.3	0.00	72	88.9	0.59
Hands-free device use 1995 (categorical)	100	96.0	0.22	28	96.4	0.00	72	95.8	0.29
Hands-free device use 1990 (categorical)	100	98.0	-0.01	28	96.4	0.00	72	98.6	0.00
Year starting using mobile phone	128	30.5	0.00	37	32.4	0.00	91	29.7	0.66
Recent mobile phone use									
Past 3 months: mobile phone use (frequency, categorical)	109	74.3	0.69	31	77.4	0.54	78	73.1	0.71
Past 3 months: mobile phone use (duration, categorical)	107	57.9	0.73	32	56.2	0.62	75	58.7	0.76
Past 3 months: hands-free use (categorical)	100	68.0	0.49	28	67.9	0.61	72	68.1	0.45
Number of mobile phones (continuous)	71	87.3	0.70	23	87.5	0.55	48	87.5	0.78
Past 3 months: number of text messages (categorical)	143	69.2	0.68	40	67.5	0.77	103	69.9	0.64
Laterality (left, right, equal, categorical)	138	78.3	0.69	38	84.2	0.76	100	76.0	0.66
Cordless phone use									
Talked on cordless phone regularly (yes/no)	225	87.6	0.39	78	93.6	0.67	147	84.4	0.25
Year starting using cordless phone at home (continuous)	160	23.1	0.62	54	24.1	0.75	106	22.6	0.56
Year starting using cordless phone at work (continuous)	64	37.5	0.62	28	39.3	0.65	36	36.1	0.60
Past 3 months: cordless phone calls (frequency, categorical)	185	61.0	0.51	68	64.7	0.58	117	57.3	0.46
Past 3 months: cordless phone calls (duration, categorical)	178	33.1	0.38	65	33.8	0.46	113	32.7	0.30

[a] Subjects who filled in the LIFEWORK baseline questionnaire twice within a period of less than 1 month (N=27) and men (N=1) were excluded from this analysis.
[b] Stratified by the number of months between participant's completion of the first subcohort's questionnaire and the second subcohort's questionnaire; 5 months approximates the median of 151 days.

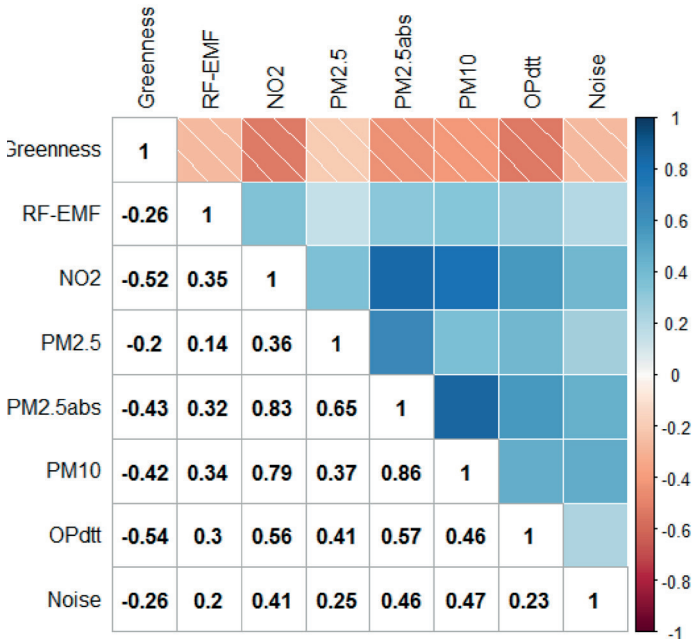
Table S3. Self-reported health outcomes[a] (prevalence or ever diagnosed) at LIFEWORK baseline (2011–2012; n=88 466)

Disease	N	%
Asthma	5 724	6.5
COPD	2 667	3.0
Angina pectoris	2 163	2.4
Heart attack	967	1.1
Other heart condition	3 739	4.2
Stroke / TIA	1 230	1.4
High blood pressure	16 618	18.8
Type I diabetes	434	0.5
Type II diabetes	2 558	2.9
High cholesterol	9 951	11.2
Chronic liver disease	351	0.4
Thyroid-related disorder	5 273	6.0
Chronic fatigue syndrome	745	0.8
Migraine	9 003	10.2
Concussion (ever)	8293	9.4
Epilepsy	756	0.9
Parkinson's disease	150	0.2
ALS	59	0.1
MS	322	0.4
Dementia/Alzheimer's disease	85	0.1
Depression	7470	8.4
Deaf or hard of hearing	4656	5.3
Cataracts	3589	4.1

ALS, Amyotrophic Lateral Sclerosis; COPD, Chronic Obstructive Pulmonary Disease; MS, Multiple Sclerosis; TIA, Transient Ischemic Attack.

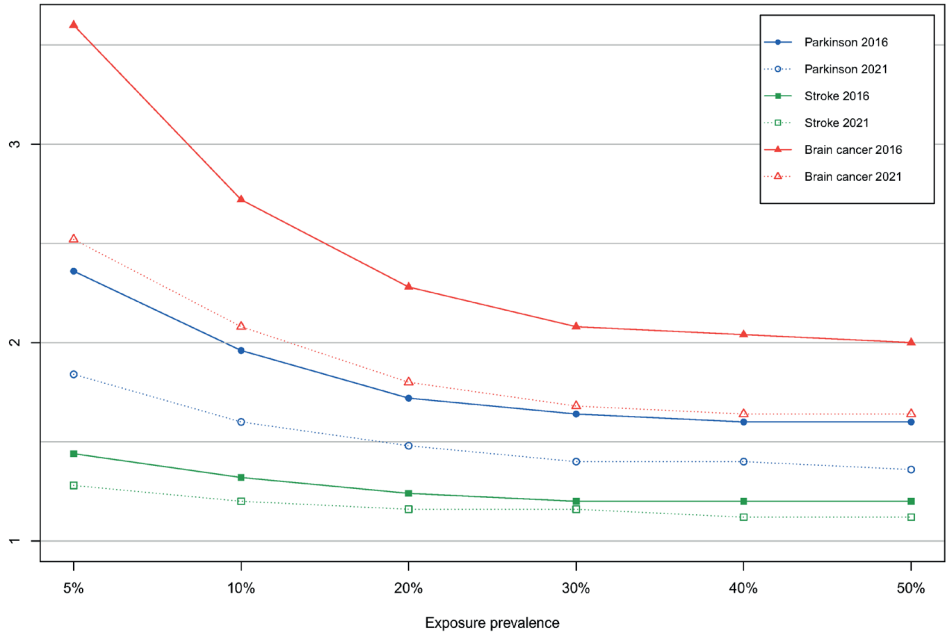
[a] This set of outcomes was assessed in the baseline and follow-up questionnaires of all three subcohorts: "Did a doctor ever diagnose you with any of the following diseases or conditions?" and age at first diagnosis. Additional outcomes were assessed in some subcohorts, as indicated in footnote (e) of Table S1.

Figure S1. Correlation matrix displaying the Spearman correlation coefficients between environmental modeled exposures among participants' baseline home addresses (n=82 979).



Greenness measured by Normalized Difference Vegetation Index, ranging from -1 to 1; noise (L_{den}) level day-evening-night from road traffic measured in dB; NO_2, nitrogen dioxide in µg/m^3; OPdtt, oxidative potential measured by dithiothreitol in nmol DTT/min/m^3; $PM_{2.5}$, particulate matter with diameter ≤2.5µm in µg/m^3; $PM_{2.5abs}$, reflectance on $PM_{2.5}$ filters (marker of black carbon) in $10^{-5}m^{-1}$; PM_{10}, particulate matter with diameter ≤10µm in µg/m^3; RF-EMF, radiofrequency electromagnetic fields from base stations in mWm^2.

Figure S2. Minimal detectable relative risks (RR) by the end of 2016 and 2021 in the LIFEWORK cohort for stroke (age-standardized mortality: 0.696/1000), Parkinson's disease (age-standardized incidence: 0.074/1000), and brain cancer (age-standardized mortality: 0.006/1000) at various exposure prevalences (α = 0.05, statistical power 80%).



REFERENCES

1. Newby DE, Mannucci PM, Tell GS, *et al.* Expert position paper on air pollution and cardiovascular disease. *Eur Heart J* 2015;**36**:83–93. doi:10.1093/eurheartj/ehu458
2. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Non-ionizing radiation, Part 2: Radiofrequency electromagnetic fields. *IARC Monogr Eval Carcinog Risks Hum* 2013;**102**:1–460.http://www.ncbi.nlm.nih.gov/pubmed/24772662
3. Pekkanen J, Pearce N. Environmental Epidemiology : Challenges and Opportunities. *Ehp* 2001;**109**:1–5.
4. Patel CJ. Analytic Complexity and Challenges in Identifying Mixtures of Exposures Associated with Phenotypes in the Exposome Era. *Curr Epidemiol Reports* 2017;**4**:22–30. doi:10.1007/s40471-017-0100-5
5. Rappaport SM. Implications of the exposome for exposure science. *J Expo Sci Environ Epidemiol* 2011;**21**:5–9. doi:10.1038/jes.2010.50
6. Wild CP. The exposome: from concept to utility. *Int J Epidemiol* 2012;**41**:24–32. doi:10.1093/ije/dyr236
7. Forouzanfar MH, Alexander L, Anderson HR, *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;**386**:2287–323. doi:10.1016/S0140-6736(15)00128-2
8. Rappaport SM. Discovering environmental causes of disease. *J Epidemiol Community Health* 2012;**66**:99–102. doi:10.1136/jech-2011-200726
9. Lim SS, Vos T, Flaxman AD, *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;**380**:2224–60. doi:10.1016/S0140-6736(12)61766-8
10. Goldberg M, Carton M, Descatha A, *et al.* CONSTANCES: a general prospective population-based cohort for occupational and environmental epidemiology: cohort profile. *Occup Environ Med* 2017;**74**:66–71. doi:10.1136/oemed-2016-103678
11. Beulens JWJ, Monninkhof EM, Verschuren WMM, *et al.* Cohort Profile: The EPIC-NL study. *Int J Epidemiol* 2010;**39**:1170–8. doi:10.1093/ije/dyp217
12. Slottje P, Yzermans CJ, Korevaar JC, *et al.* The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in the Netherlands. *BMJ Open* 2014;**4**:e005858. doi:10.1136/bmjopen-2014-005858
13. Pijpe A, Slottje P, van Pelt C, *et al.* The Nightingale study: rationale, study design and baseline characteristics of a prospective cohort study on shift work and breast cancer risk among nurses. *BMC Cancer* 2014;**14**:47. doi:10.1186/1471-2407-14-47
14. Schüz J, Elliott P, Auvinen A, *et al.* An international prospective cohort study of mobile phone users and health (Cosmos): Design considerations and enrolment. *Cancer Epidemiol* 2011;**35**:37–43. doi:10.1016/j.canep.2010.08.001
15. Toledano MB, Auvinen A, Tettamanti G, *et al.* An international prospective cohort study of mobile phone users and health (COSMOS): factors affecting validity of self-reported mobile phone use. *Int J Hyg Environ Health* Published Online First: 2017. doi:https://doi.org/10.1016/j.ijheh.2017.09.008
16. Kosinski M, Bayliss MS, Bjorner JB, *et al.* A six-item short-form survey for measuring headache impact: the HIT-6. *Qual Life Res* 2003;**12**:963–74. doi:10.1023/A:1026119331193
17. Lipton RB, Dodick D, Sadovsky R, *et al.* A self-administered screener for migraine in primary care. *Neurology* 2003;**61**:375–82. doi:10.1212/01.WNL.0000078940.53438.83
18. Hays RD, Martin S a, Sesti AM, *et al.* Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005;**6**:41–4. doi:10.1016/j.sleep.2004.07.006

19. Burney PGJ, Luczynska C, Chinn S, *et al.* The European Community Respiratory Health Survey. *Eur Respir J* 1994;**7**:954–60. doi:10.1183/09031936.94.07050954
20. John E . Ware , Jr . MK and SD. K. A 12-Item Short-Form Health Survey : Construction of Scales and Preliminary Tests of Reliability and Validity. *Lippincott Williams & Wilkins Stable* 1996;**34**:220–33.
21. Martens AL, Slottje P, Timmermans DRM, *et al.* Modeled and Perceived Exposure to Radio-Frequency Electromagnetic Fields From Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort. *Am J Epidemiol* 2017;:1–10. doi:10.1093/aje/kwx041
22. Roenneberg T, Wirz-Justice A, Merrow M. Life between Clocks: Daily Temporal Patterns of Human Chronotypes. *J Biol Rhythms* 2003;**18**:80–90. doi:10.1177/0748730402239679
23. Beekhuizen J, Vermeulen R, Kromhout H, *et al.* Geospatial modelling of electromagnetic fields from mobile phone base stations. *Sci Total Environ* 2013;**445–446**:202–9. doi:10.1016/j.scitotenv.2012.12.020
24. Beekhuizen J, Vermeulen R, van Eijsden M, *et al.* Modelling indoor electromagnetic fields (EMF) from mobile phone base stations for epidemiological studies. *Environ Int* 2014;**67**:22–6. doi:10.1016/j.envint.2014.02.008
25. Eeftens M, Beelen R, de Hoogh K, *et al.* Development of Land Use Regression Models for PM 2.5 , PM 2.5 Absorbance, PM 10 and PM coarse in 20 European Study Areas; Results of the ESCAPE Project. *Environ Sci Technol* 2012;**46**:11195–205. doi:10.1021/es301948k
26. Beelen R, Hoek G, Vienneau D, *et al.* Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. *Atmos Environ* 2013;**72**:10–23. doi:10.1016/j.atmosenv.2013.02.037
27. Rhew IC, Vander Stoep A, Kearney A, *et al.* Validation of the Normalized Difference Vegetation Index as a Measure of Neighborhood Greenness. *Ann Epidemiol* 2011;**21**:946–52. doi:10.1016/j.annepidem.2011.09.001
28. Scheurs EM, Jabben J, Verheijen ENG. STAMINA-Model description STAMINA - Model description. Rep. 680740003/2010. 2010. http://www.rivm.nl/bibliotheek/rapporten/680740003.pdf
29. Locatus. http://locatus.com/nederland (accessed 17 Aug 2016).
30. de Vocht F, Zock J-P, Kromhout H, *et al.* Comparison of self-reported occupational exposure with a job exposure matrix in an international community-based study on asthma. *Am J Ind Med* 2005;**47**:434–42. doi:10.1002/ajim.20154
31. Matheson MC. Biological dust exposure in the workplace is a risk factor for chronic obstructive pulmonary disease. *Thorax* 2005;**60**:645–51. doi:10.1136/thx.2004.035170
32. Peters S, Vermeulen R, Olsson A, *et al.* Development of an Exposure Measurement Database on Five Lung Carcinogens (ExpoSYN) for Quantitative Retrospective Occupational Exposure Assessment. *Ann Occup Hyg* 2012;**56**:70–9. doi:10.1093/annhyg/mer081
33. Peters S, Vermeulen R, Cassidy A, *et al.* Comparison of exposure assessment methods for occupational carcinogens in a multi-centre lung cancer case-control study. *Occup Environ Med* 2011;**68**:148–53. doi:10.1136/oem.2010.055608
34. Huss A, Vermeulen R, Bowman JD, *et al.* Electric shocks at work in Europe: development of a job exposure matrix. *Occup Environ Med* 2013;**70**:261–7. doi:10.1136/oemed-2012-100732
35. Koeman T, Offermans NSM, Christopher-de Vries Y, *et al.* JEMs and Incompatible Occupational Coding Systems: Effect of Manual and Automatic Recoding of Job Codes on Exposure Assignment. *Ann Occup Hyg* 2013;**57**:107–14. doi:10.1093/annhyg/mes046
36. Kadaster Netherlands. Kadaster Netherlands. https://www.kadaster.nl/ (accessed 17 Aug 2016).
37. Jedynska A, Hoek G, Wang M, *et al.* Development of land use regression models for elemental, organic carbon, PAH, and hopanes/steranes in 10 ESCAPE/TRANSPHORM European study areas. *Environ Sci Technol* 2014;**48**:14435–44. doi:10.1021/es502568z

38. Brouwer M, Koeman T, van den Brandt PA, *et al.* Occupational exposures and Parkinson's disease mortality in a prospective Dutch cohort. *Occup Environ Med* 2015;**72**:448–55. doi:10.1136/oemed-2014-102209

39. Bowman JD, Touchstone J a, Yost MG. A Population-Based Job Exposure Matrix for Power-Frequency Magnetic Fields. *J Occup Environ Hyg* 2007;**4**:715–28. doi:10.1080/15459620701528001

40. Statistics Netherlands. www.cbs.nl (accessed 24 Oct 2016).

41. van der Sanden GA, Coebergh JW, Schouten LJ, Visser O van LF. Cancer incidence in The Netherlands in 1989 and 1990: first results of the nationwide Netherlands cancer registry. Coordinating Committee for Regional Cancer Registries. *Eur J Cancer* 1995;**11**:1822–9.

42. NIVEL Primary Care Database. https://www.nivel.nl/en/dossier/nivel-primary-care-database (accessed 17 Aug 2016).

43. Schweikardt C, Verheij RA, Donker GA, *et al.* The historical development of the Dutch Sentinel General Practice Network from a paper-based into a digital primary care monitoring system. *J Public Health (Bangkok)* Published Online First: 15 August 2016. doi:10.1007/s10389-016-0753-4

44. Karasek R, Brisson C, Kawakami N, *et al.* The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol* 1998;**3**:322–55. doi:10.1037/1076-8998.3.4.322

45. Terluin B, van Marwijk HW, Adèr HJ, *et al.* The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;**6**:34. doi:10.1186/1471-244X-6-34

46. de Graaf R, ten Have M, van Gool C, *et al.* Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Soc Psychiatry Psychiatr Epidemiol* 2012;**47**:203–13. doi:10.1007/s00127-010-0334-8

47. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol* 2013;**42**:1012–4. doi:10.1093/ije/dys223

48. Howe CJ, Cole SR, Lau B, *et al.* Selection Bias Due to Loss to Follow Up in Cohort Studies. *Epidemiology* 2016;**27**:91–7. doi:10.1097/EDE.0000000000000409

49. Naimi AI, Richardson DB, Cole SR. Causal Inference in Occupational Epidemiology: Accounting for the Healthy Worker Effect by Using Structural Nested Models. *Am J Epidemiol* 2013;**178**:1681–6. doi:10.1093/aje/kwt215

50. Weisskopf MG, Sparrow D, Hu H, *et al.* Review Biased Exposure – Health Effect Estimates from Selection in Cohort Studies : Are Environmental Studies at Particular Risk ? *Ehp* 2015;**123**:1113–22.

51. Toledano MB, Smith RB, Brook JP, *et al.* How to Establish and Follow up a Large Prospective Cohort Study in the 21st Century - Lessons from UK COSMOS. *PLoS One* 2015;**10**:e0131521. doi:10.1371/journal.pone.0131521

52. Statistics Netherlands Statline. http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=71098ned&D1=33,58&D2=0&D3=a&HDR=G1,T&STB=G2&VW=T (accessed 24 Sept 2016).

53. Horevoorts NJE, Vissers PAJ, Mols F, *et al.* Response rates for patient-reported outcomes using web-based versus paper questionnaires: Comparison of two invitational methods in older colorectal cancer patients. *J Med Internet Res* 2015;**17**:1–12. doi:10.2196/jmir.3741

54. van Gelder MMHJ, Bretveld RW, Roeleveld N. Web-based Questionnaires: The Future in Epidemiology? *Am J Epidemiol* 2010;**172**:1292–8. doi:10.1093/aje/kwq291

55. Vrijheid M, Armstrong BK, Bédard D, *et al.* Recall bias in the assessment of exposure to mobile phones. *J Expo Sci Environ Epidemiol* 2009;**19**:369–81. doi:10.1038/jes.2008.27

56. Cardis E. Brain tumour risk in relation to mobile telephone use: Results of the INTERPHONE international case-control study. *Int J Epidemiol* 2010;**39**:675–94. doi:10.1093/ije/dyq079

57. Richardson DB, Cole SR, Langholz B. Regression Models for the Effects of Exposure Rate and Cumulative Exposure. *Epidemiology* 2012;**23**:892–9. doi:10.1097/EDE.0b013e31826c3149

58. Agier L, Portengen L, Chadeau-Hyam M, *et al.* A Systematic Comparison of Linear Regression-Based Statistical Methods to Assess Exposome-Health Associations. *Environ Health Perspect* 2016;**172**. doi:10.1289/EHP172

59. Lenters V, Vermeulen R, Portengen L. Performance of variable selection methods for assessing the health effects of correlated exposures in case-control studies. *Occup Environ Med* Published Online First: 2017. doi:10.1136/oemed-2016-104231

60. Steven G. O'Connell, Laurel D. Kincl and KAA. Silicone Wristbands as Personal Passive Samplers. *Environ Sci Technol* 2014;**48**:3327–35. doi:http://pubs.acs.org/doi/pdf/10.1021/es405022f

61. Goedhart G, Kromhout H, Wiart J, *et al.* Validating self-reported mobile phone use in adults using a newly developed smartphone application. *Occup Environ Med* 2015;**72**:oemed-2015-102808. doi:10.1136/oemed-2015-102808

REFERENCES SUPPLEMENTARY MATERIAL

1. Weinstein ND. Individual differences in reactions to noise: a longitudinal study in a college dormitory. *J Appl Psychol* 1978;**63**:458–66. doi:10.1037/0021-9010.63.4.458
2. Karasek R, Brisson C, Kawakami N, *et al.* The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol* 1998;**3**:322–55. doi:10.1037/1076-8998.3.4.322
3. Schüz J, Elliott P, Auvinen A, *et al.* An international prospective cohort study of mobile phone users and health (Cosmos): Design considerations and enrolment. *Cancer Epidemiol* 2011;**35**:37–43. doi:10.1016/j.canep.2010.08.001
4. John E . Ware , Jr . MK and SD. K. A 12-Item Short-Form Health Survey : Construction of Scales and Preliminary Tests of Reliability and Validity. *Lippincott Williams & Wilkins Stable* 1996;**34**:220–33.
5. Kosinski M, Bayliss MS, Bjorner JB, *et al.* A six-item short-form survey for measuring headache impact: the HIT-6. *Qual Life Res* 2003;**12**:963–74. doi:10.1023/A:1026119331193
6. Lipton RB, Dodick D, Sadovsky R, *et al.* A self-administered screener for migraine in primary care. *Neurology* 2003;**61**:375–82. doi:10.1212/01.WNL.0000078940.53438.83
7. Hays RD, Martin S a, Sesti AM, *et al.* Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005;**6**:41–4. doi:10.1016/j.sleep.2004.07.006
8. Juda M, Vetter C, Roenneberg T. The Munich ChronoType Questionnaire for Shift-Workers (MCTQ ^{Shift}). *J Biol Rhythms* 2013;**28**:130–40. doi:10.1177/0748730412475041
9. Roenneberg T, Wirz-Justice A, Merrow M. Life between Clocks: Daily Temporal Patterns of Human Chronotypes. *J Biol Rhythms* 2003;**18**:80–90. doi:10.1177/0748730402239679
10. Burney PGJ, Luczynska C, Chinn S, *et al.* The European Community Respiratory Health Survey. *Eur Respir J* 1994;**7**:954–60. doi:10.1183/09031936.94.07050954
11. Terluin B, van Marwijk HW, Adèr HJ, *et al.* The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;**6**:34. doi:10.1186/1471-244X-6-34
12. Sheldon Cohen T, Kamarck T, Robin Mermelstein. A Global Measure of Perceived Stress. *J Health Soc Behav* 1983;**24**:385–96.
13. Rocca WA, Maraganore DM, McDonnell SK, *et al.* Validation of a Telephone Questionnaire for Parkinson's Disease. *J Clin Epidemiol* 1998;**51**:517–23. doi:10.1016/S0895-4356(98)00017-1
14. Martens AL, Slottje P, Timmermans DRM, *et al.* Modeled and Perceived Exposure to Radio-Frequency Electromagnetic Fields From Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort. *Am J Epidemiol* 2017;:1–10. doi:10.1093/aje/kwx041

Chapter 3

Regression calibration of self-reported mobile phone use to improve quantitative risk estimation in the COSMOS study

Marije Reedijk*, Lützen Portengen*, Anssi Auvinen, Katja Kojo, Sirpa Heinävaara, Maria Feychting, Giorgio Tettamanti, Anders Ahlbom, Lena Hillert, Paul Elliott, Mireille Toledano, Joachim Schüz, Isabelle Deltour, Aslak Harbo Poulsen, Robert Verheij, Petra Peeters, Matti Rookus, Eugenio Traini, Anke Huss, Hans Kromhout, Roel Vermeulen and the COSMOS Study Group

*Equal contribution

In preparation

ABSTRACT

The Cohort Study of Mobile Phone Use and Health (COSMOS) study has repeatedly collected both self-reported and operator-recorded data on mobile phone use. Assessing health effects using self-reported information only is prone to measurement error, but operator data was not available for past mobile phone use and prospectively only for part of the study population. Here, we evaluated different statistical approaches for optimizing use of all available call data within COSMOS for constructing exposure histories. We evaluated and compared the performance of complete case-analysis, several regression calibration methods, and multiple imputation in a simulation study with a binary health outcome. We used self-reported and operator-recorded mobile phone call data collected at baseline (2007-2012) from participants in Denmark, Finland, the Netherlands, Sweden, and the UK. Parameter estimates obtained using regression calibration methods included only little bias and had a lower mean squared error than those obtained with complete-case analysis or multiple-imputation. Our study showed that regression calibration methods allowed more accurate estimation of the relation between mobile phone use and health outcomes by combining self-reported data with objective operator-recorded data available for a subset of the participants.

INTRODUCTION

Exposure measurement error is a threat to the validity of environmental epidemiological research, and may be especially important for studies that rely on self-reported exposure information[1-3]. Systematic and random measurement error may result in substantial bias and loss of precision in estimated exposure-outcome relations[1,3]. Validation studies have shown substantial error in self-reported estimates of mobile phone use[4-10]. Measurement error correction methods, such as regression calibration (RC), have been widely used in nutritional epidemiology[2,11-13], and are increasingly used in environmental epidemiology[3,14-18].

For mobile phone use, multiple bias modelling has been used in the Interphone Study[19], bias correction in the CEFALO study[6], and regression calibration by Tokola *et al.*[20]. Redmayne *et al.* developed a measurement error correction method for the number of weekly text messages[16]. The prospective *COhort Study of MObile phone uSe and health* (COSMOS)[9,21] aimed to collect both self-reported and operator-recorded mobile phone use data for more than 310,000, participants from 6 countries. Operator data could not be obtained for all participants, e.g. because subscriptions were by employers, prepaid service, data sharing issues, or because participants did not provide informed consent[21]. Exposure-outcome relations within COSMOS could be estimated using either self-reported or operator-recorded mobile phone use data. When comparing self-reported to operator-recorded information, the validity of self-reported mobile phone use within COSMOS was shown to depend on sex, age, and self-reported symptoms[10]. Deciding which information to use for the primary analyses requires a careful tradeoff. Using operator-recorded information only may result in low precision estimates, because data is available for a subset of participants only, and could lead to biased risk estimates if data is not missing at random and because of measurement error. Using self-reported mobile phone use is likely to result in biased risk estimates and precision loss due to even greater measurement error.

The aim of the current paper is to address an important concern: how to leverage self-reported exposure estimates that are easily available, but error-prone, with more objective measurements that may be obtained for a subset of participants only. Here we investigated several alternative approaches for combining self-reported and operator-recorded mobile phone use within the COSMOS study. We applied different variants of regression calibration (RC)[3,22,23], and compared these to complete-case (CC) analyses and multiple imputation (MI)[24] in a simulation study, evaluating bias and precision in estimated slope coefficients of a logistic regression model.

METHODS

Study design

We used data from the COSMOS study, which has been described in detail elsewhere[21,25,26]. Briefly, information on mobile phone use was collected for subjects aged 18 and over from 2007 onwards in six countries: Denmark, Finland, France, the Netherlands, Sweden, and the United Kingdom (UK)[21,25,26]. Table 1 compares enrolment periods and recruitment strategies. In Denmark, Finland, and Sweden, potential participants were selected using stratified random sampling from subscribers of major mobile phone network operators, with strata defined by call duration, age, and sex (Finland only). In the Netherlands, recruitment was from the general population and nurses[26]. In France, recruitment was from the general population. In the UK, approximately 65% of participants were recruited by stratified random sampling from subscribers of major mobile phone network operators, with strata defined by call duration, age, and sex, while 35% of participants were recruited from the electoral registry[25]. The COSMOS study protocol was approved by ethical committees in each country. Written or electronic informed consent to link to operator-recorded mobile phone use data was obtained for each participant, except in France and Sweden, where enrollment in the study was not conditional on consent.

Table 1. Study design: participating countries, enrolment year, and recruitment strategy.

Country	Enrolment		Recruitment strategy	Questionnaire
	Year	N		
Denmark	2007-2009	25912	Stratified random sampling from operator subscription data (a)	Paper
Sweden	2008-2009	55471	Stratified random sampling from operator subscription data (a)	Paper
Finland	2009-2010	13062	Stratified random sampling from operator subscription data (b)	Paper/Electronic
UK	2009-2012	98617*	Stratified random sampling from operator subsctiption data (b) / General population sampling	Electronic
Netherlands	2011-2012	88466	General populaton sampling and occupational population sampling (nurses)	Paper/ Electronic

(a) Stratified random sampling from operator subscription data based on call time and age
(b) Stratified random sampling from operator subscription data based on call time, age, and sex
* data on participants from one operator in the UK were not available for the present analyses.

Data collection

The COSMOS baseline questionnaire was administered between 2007-2012 either on paper or electronically in most countries. It was administered 2017-2019 in France, and the French data was therefore not considered in the current analysis. Key topics included mobile phone use, cordless

phone use, use of other wireless devices, demographic and social characteristics and several self-reported health outcomes[21]. Characteristics considered as potential predictors of mobile phone use (or affecting self-report) were sex, age, marital status (living together, living apart, and not being in a relationship), educational level (elementary versus secondary school or higher), and employment status (active versus inactive, i.e. either student, unemployed, retired, or on sick leave).

Self-reported mobile phone use

Baseline mobile phone use assessment[21] included information on the number of phones and the frequency of calls. Self-reported duration of mobile phone use (REPORT) was based on answers to the question: "*Over the last three months, on average, how much time per week did you spend talking on a mobile phone?*". The following response options were provided: "< 5 min/week", "5-29 min/week", "30-59 min/week", "1-3 hours/week", "4-6 hours/week", and ">6 hours/week". In the Netherlands and the UK two further categories of call duration were considered, "7-9 hours/week" and "10 or more hours/week", but these were combined for the present analyses.

Operator-recorded mobile phone use

Operator-recorded duration of mobile phone use (RECORD) was collected for all participants who provided consent and had a subscription under their own name. In some countries, operator-recorded data was not available for participants who used a business or pre-paid mobile phone. Operator-recorded data was used only when available for all mobile phones that were reported (up to two phones in DK, NL, SE, FI, up to three in the UK) and when available for at least three full months at the time the baseline questionnaire was administered. Participants with mobile phones that were also used by others were excluded. Operator-recorded data was available for 12% of participants in DK, 76% in FI, 4% in the NL, 48% in SE, and 39% in the UK.

Providers collected information on number and duration separately for outgoing and incoming calls. Across countries, outgoing calls made up 55-72% of the total number. This suggests under-recording of incoming calls, possibly because calls between subscribers from the same provider were not always charged and may have gone unrecorded in some billing systems. For this reason, we based our simulations on outgoing call duration only. Data for the three months at baseline were used to calculate average duration of calls in minutes per week.

SIMULATION STUDY

We compared results from different approaches that combine self-reported and operator-recorded mobile phone use: a complete-case (CC) analysis, four variants of regression calibration (simple RC, Generalized Additive Model for Location, Shape, and Scale (GAMLSS)-based RC, direct RC, and inverse

RC), and multiple imputation (MI) in a simulation study. A roadmap to the main results from this study is provided in Figure 1.

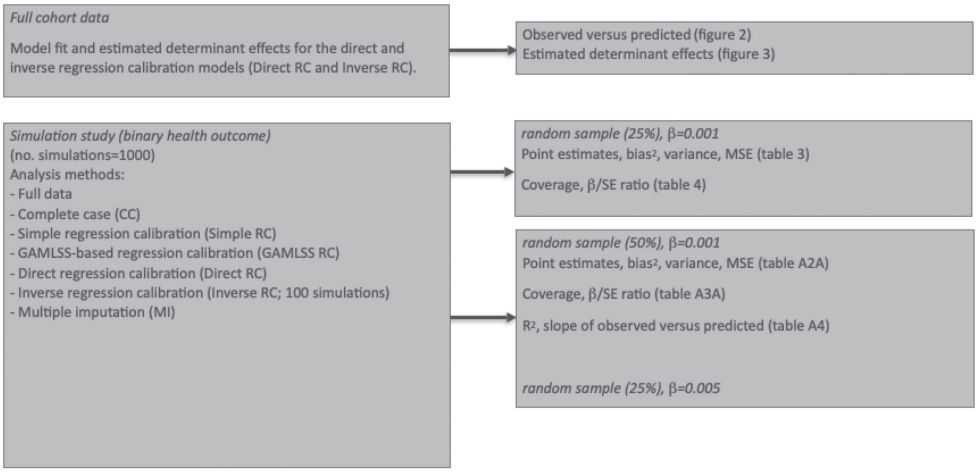
All simulations used data from the subset of participants, for whom both self-reported and operator-recorded data were available. For each simulation, we assigned a binary health outcome with probability $P(Y_i=1)$ according to the following equation:

$$P(Y_i=1) = \exp(\alpha + b*RECORD_i)/(1+ \exp(\alpha + b*RECORD_i))$$

for subject *i* with recorded call duration $RECORD_i$ (in minutes/week).

For our main analyses, the outcome was simulated using a slope coefficient (β) of 0.001 (i.e. assuming an Odds Ratio of exp(0.02)=1.02 per additional 20 minutes call-time per week) and with a balanced ratio of cases to non-cases (by tuning α). The value of 0.001 was chosen to provide a reasonably strong signal-to-noise ratio across the country-specific datasets, while avoiding substantial bias for the RC methods[31]. Additional results using a slope coefficient (β) of 0.005 (i.e. assuming an Odds Ratio of exp(0.1)=1.11 per additional 20 minutes call-time per week) are presented in the appendix.

Figure 1. Roadmap to the main results from this paper. This includes evaluation of model fit and estimated determinant effects for RC models fitted to the available cohort data and results of the simulation study (including sensitivity analyses).



We randomly assigned 25% of the participants to the training set used to develop the regression calibration models and for fitting the complete-case model, while the remainder was used as a test set for fitting the "calibrated" health outcome model (mimicking the situation that provider data was missing for 75% of the population). Results obtained using 50% of subjects as the training set are presented in the appendix.

The performance of each approach was evaluated by comparing the average estimated slope to the "true" slope coefficient used in the simulations and by estimating squared bias, variance, and mean squared error (MSE). We calculated 95% confidence intervals [95%CI] using bootstrapping to reflect Monte-Carlo error from using a limited number of simulations. We also estimated the coverage of 95% CIs calculated using the estimated standard error of the slope coefficient (Wald-type 95% CIs) to investigate whether these correctly accounted for the additional (sampling) variability. We calculated ratios of estimated slopes to estimated standard errors (β/SE) as a surrogate for statistical power and included results for the full data health model in the tables as an "ideal-case" reference method.

We used 1,000 simulations to evaluate performance for all approaches, except for the inverse RC method, where 100 simulations were used to avoid excessive running times (fitting a single model could take more than a day for the larger datasets). For the same reason, we used model-based ("naïve") standard errors to obtain precision-weighted estimates, as we found these to be only marginally different from bootstrapped standard errors in a subset of 100 simulations.

Full data and complete-case analysis

Logistic regression models were fitted to the simulated health outcome in the combined training and test datasets ("full data") or the training set only ("complete-case"). Note that the "full data" option would not be available for COSMOS.

Operator-recorded mobile phone use was treated as gold standard for "*true"* mobile phone use, and the goal of our analyses was to estimate the slope coefficient (β) in the following health outcome model:

$$g(E(Y)) = \alpha + \beta.\text{RECORD}$$

where g is the logistic link function in case of a binary outcome Y, and RECORD is operator-recorded mobile phone use.

Regression calibration

Because RECORD is only available for a subset of the participants, we consider using self-reported mobile phone use (REPORT) as an (error-prone) proxy for operator-recorded use (RECORD) and consider different implementations of regression calibration (RC) to correct for the measurement error[27]. All involve fitting the so-called "calibrated" health outcome model:

$$g(E(Y)) = a^* + \beta^*.E(\text{RECORD}|\text{REPORT},Z)$$

where E(RECORD|REPORT,Z) is the expected value of "*true*" operator-recorded mobile phone use, conditional on REPORT and other covariates (Z).

We evaluated four different RC approaches, which differ by the amount of information used and in assumptions regarding the distribution of RECORD or on the relation between REPORT and RECORD. Our simple RC approach is the most basic, using the empirical average RECORD for each REPORT category. REPORT estimates appeared to follow an approximate log-normal distribution, and for the model-based approaches we indirectly estimated E(RECORD|REPORT,Z) using the known relation between the arithmetic mean (AM) and geometric mean (GM) plus geometric standard deviation (GSD) of a log-normally distributed variable (i.e. $AM = GM*\exp(\log(GSD)^2/2)$). Our second approach, GAMLSS RC, uses maximum likelihood in a generalized additive modeling framework to estimate both GM and GSD as a function of REPORT categories and covariates (sex, age, educational achievement, employment status, and marital status), allowing for non-linear effects of continuous covariates. Our third approach, which we call direct RC, uses regression splines to achieve the same flexibility in a Bayesian setting, but in addition relaxes distributional assumptions by modelling the residuals non-parametrically as a Dirichlet process mixture of (log)normal distributions, and with the AM estimated by averaging across draws from that distribution. Finally, our fourth approach, which we call inverse RC, is a Bayesian structural method and features a probit regression model for REPORT categories using RECORD and all other covariates as predictor variables. By modelling the prior (conditional) distribution of RECORD as a Dirichlet process mixture of (log)normal distributions, we may inverse the model to obtain estimates of RECORD conditional on REPORT and other covariates.

The GAMLSS RC model was fitted using the R package gamlss[27], while the Bayesian RC models (direct, inverse) were fitted using the R package rjags[28]. A more detailed description of all models is provided in the appendix.

Because health data in the COSMOS study will be available for most of the participants that contribute to the RECORD models (i.e. have operator-recorded data), we follow the approach suggested by Spiegelman et al.[29] for studies that have an internal validation study. This approach consists of combining the slope estimate from the complete-case model (i.e. the health model for the 25% of participants where RECORD was available) with that from the "calibrated" model fit to data from the remainder of the population (the test set), using inverse-precision weighting. Although robust standard errors for the slope estimate from the "calibrated" model could be obtained using bootstrapping, we did not do that here for computational reasons.

Multiple imputation

Measurement error can also be regarded as a missing data problem[30], and we therefore considered multiple imputation (MI) in addition to the complete-case and RC analyses for the simulations. MI of operator-recorded phone use was performed by chained equations with fully conditional specification, using predictive mean matching as implemented in the R package mice[24]. Participant characteristics considered as predictors in the imputation model were the same as for the RC approaches (i.e. sex, age, educational achievement, employment status, and marital status), but also included the (simulated) health outcome (Y). We imputed a total of 10 different datasets and combined exposure slope coefficients from the health outcome model using Rubin's rule[31].

Model fit

We evaluated model fit for RC models using data from all participants with complete RECORD, REPORT, and the covariate data, for each country separately, by comparing predicted to observed average weekly call-time. We calculated the proportion of variance in observed call durations that could be explained by predicted call durations and evaluated the relation between predicted and observed call durations in detail by fitting a GAMLSS model that allowed for non-linear effects and non-homogeneous residual variance.

RESULTS

Subject characteristics

Table 2 lists study attrition and subject characteristics for included subjects. Complete information on REPORT (self-reported data) and RECORD (operator-recorded data) was available for <50% of subjects for most countries, except for Finland. In the Netherlands participants were sampled from the general population and nurses, complete information was available for only 3% and REPORT was not filled out by 33% of subjects due to the high percentage of non-users. Slightly more women than men participated, except for the Netherlands, where most participants were women. There were notable differences in REPORT and RECORD between countries. RECORD was lowest in the Netherlands (GM [GSD] = 23.4 [2.6] minutes/week) and highest in Finland (81.5 [3.0] minutes/week).

A detailed comparison of participants with or without information on self-reported and operator-recorded mobile phone use and covariates per country is provided in Appendix Table A1. Missing RECORD data was significantly associated with REPORT in all countries, but absolute differences were small and the direction across countries was inconsistent. Age, sex, marital status, employment status, and educational level were also significantly associated with missing RECORD information in most countries.

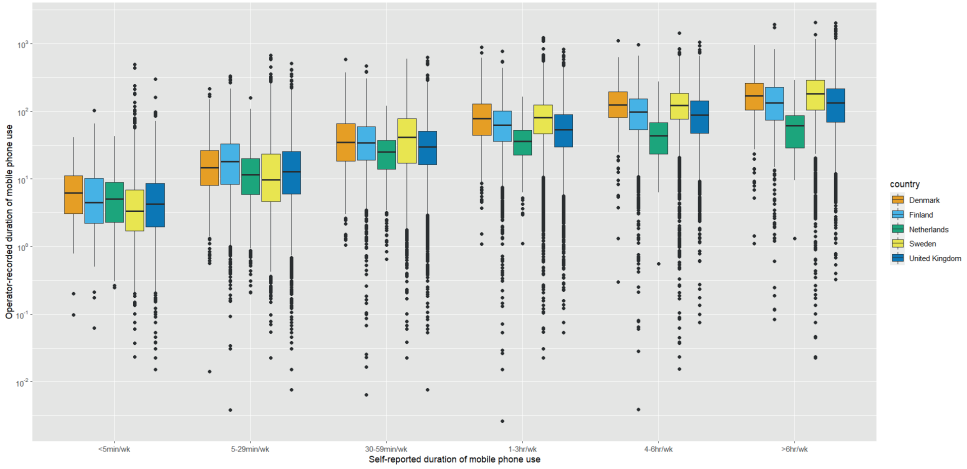
Table 2. Overview of available data and subject characteristics for participants where both self-reported and operator-recorded data were available.

	Denmark	Finland	Netherlands	Sweden	United Kingdom
Total participants (n)	25912	13062	88466	55471	98617
Covariate data missing (n [%])	706 [3%]	260 [2%]	3428 [4%]	2641 [5%]	11908 [12%]
No-user or self-reported use missing (n [%])	1871 [7%]	550 [4%]	29160 [33%]	5094 [9%]	6375 [6%]
Operator data missing (n [%])	20342 [79%]	3090 [24%]	52839 [60%]	22855 [41%]	48220 [49%]
Complete data (n [%])	2993 [12%]	9162 [70%]	3039 [3%]	24881 [45%]	32114 [33%]
Age (mean [sd])					
- Years	46.9 [12.9]	49.3 [14.0]	45.0 [12.3]	43.3 [13.4]	44.2 [14.8]
Sex (n [%])					
- Male	1350 [45%]	3808 [42%]	280 [9%]	11352 [46%]	14365 [45%]
- Female	1643 [55%]	5354 [58%]	2759 [91%]	13529 [54%]	17749 [55%]
Marital status (n [%])					
- Living together	2004 [67%]	6386 [70%]	2282 [75%]	16784 [67%]	20900 [65%]
- Not living together	321 [11%]	860 [9%]	131 [4%]	2768 [11%]	4259 [13%]
- Not in a relationship	668 [22%]	1916 [21%]	626 [21%]	5329 [21%]	6955 [22%]
Employment status (n [%])					
- Active	2212 [74%]	4721 [52%]	2562 [84%]	18033 [72%]	22374 [70%]
- Inactive*	781 [26%]	4441 [48%]	477 [16%]	6848 [28%]	9740 [30%]
Educational level (n [%])					
- Elementary school	423 [14%]	4014 [44%]	165 [5%]	3331 [13%]	19563 [61%]
- At least secondary school	2570 [86%]	5148 [56%]	2874 [95%]	21550 [87%]	12551 [39%]
Self-reported call duration (n [%])					
< 5 min/week	132 [4%]	101 [1%]	378 [12%]	1021 [4%]	1238 [4%]
5-29min/week	1086 [36%]	1634 [18%]	1555 [51%]	7023 [28%]	8749 [27%]
30-59 min/week	661 [22%]	2175 [24%]	642 [21%]	4894 [20%]	7466 [23%]
1-3 hours/week	676 [23%]	3573 [39%]	359 [12%]	6647 [27%]	9090 [28%]
4-6 hours/week	268 [9%]	1181 [13%]	75 [2%]	3075 [12%]	3321 [10%]
>6 hours/week	170 [6%]	498 [5%]	30 [1%]	2221 [9%]	2250 [7%]
Recorded call duration (GM [GSD])					
- Minutes/week	60.6 [3.3]	81.5 [3.0]	23.4 [2.6]	78.3 [3.9]	49.1 [3.4]

*Inactive includes ritered, school/studying, unemployed, housewife/househusband, disabled

Figure 2 shows the distribution of RECORD for outgoing call only, by country and categories of self-reported use. RECORD tended to be lower for participants from the NL and the UK, even within categories of self-reported use, and this was most pronounced for higher REPORT categories.

Figure 2. Outgoing operator-recorded mobile phone use by country and categories of self-reported use.

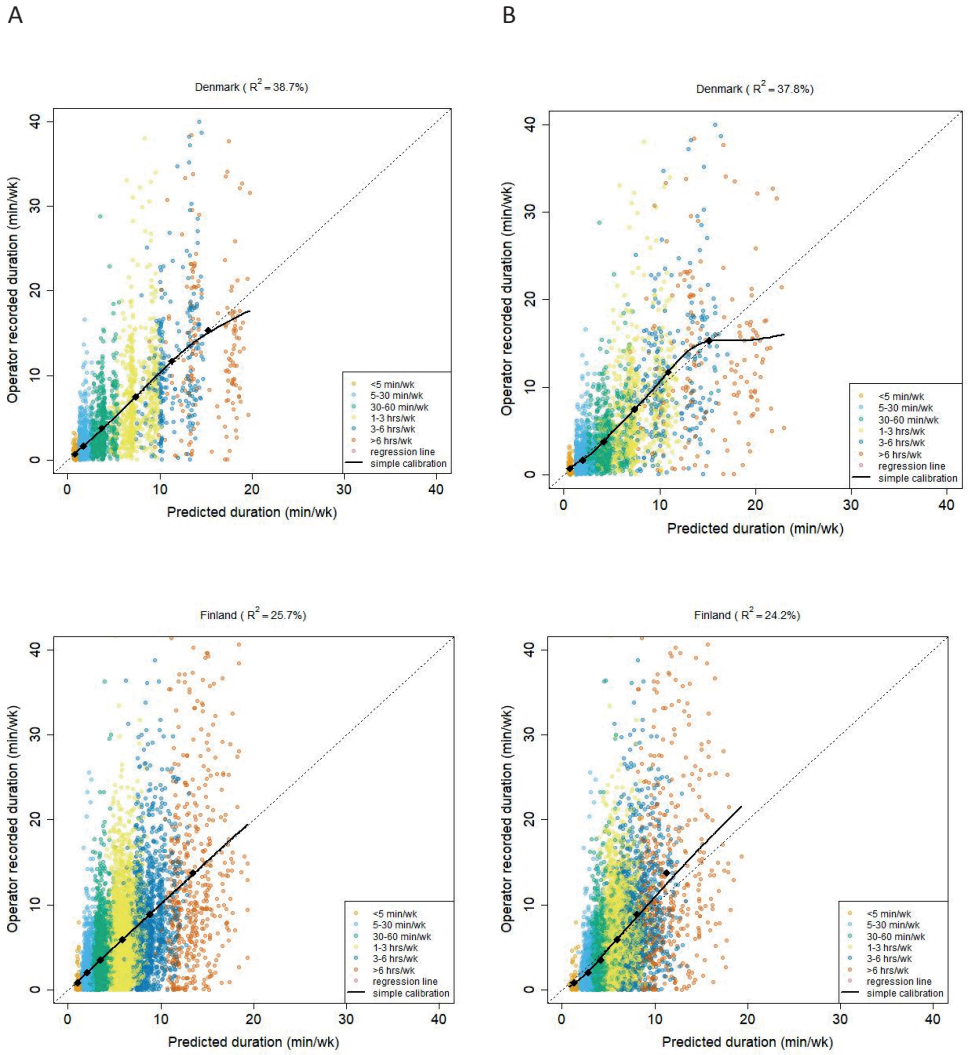


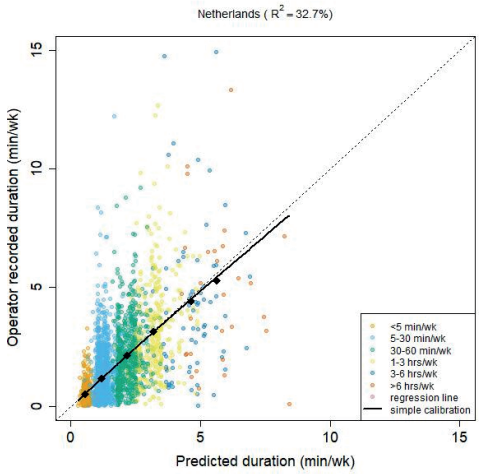
MODEL FIT AND PARAMETER ESTIMATES FROM RC MODELS ON THE FULL DATA

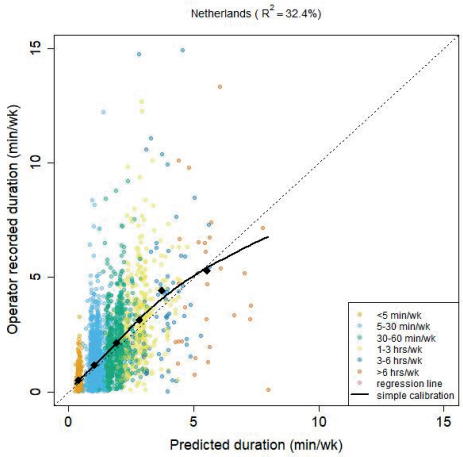
Estimated covariate effects for the direct and inverse RC models fitted to the full cohort data, by country, are presented in the appendix (Figures A1 and A2). Covariate effects were similar for the direct and inverse RC methods within and between countries. Being in a relationship but not living together was a strong predictor for RECORD in all countries and was also associated with higher REPORT categories (conditional on RECORD) in the probit regression model for the inverse RC method (Figure A1). Older age was associated with lower RECORD and REPORT in most countries (Figure A2A/B). The underlying variable for REPORT in the probit regression model for the inverse RC method tended to decrease with age (Figure A2C), but increased with increasing RECORD (Figure A2D).

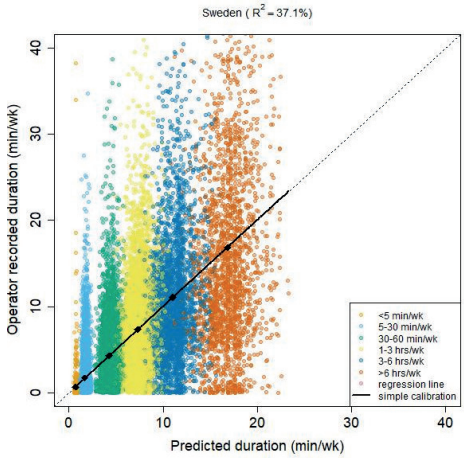
The relation between observed and predicted RECORD was close to linear for the direct RC model in most countries, but predictions in the high RECORD range tended to either underestimate (Finland, Sweden) or overestimate (Denmark, the Netherland, the UK) average observed values for the inverse RC model (Figure 3). Predictions from the GAMLSS-based RC model tended to be higher than observed in the high RECORD range for all countries (Appendix Figure A3). Differences in (in-sample) R^2 between models were smaller than differences between countries, but were consistently larger for the direct than the inverse RC method. R^2 was largest for RC models fitted to the Danish data (range 36.1-38.7%) and lowest for models fitted to the Finnish data (range 23.6-25.7%). Using any of the more complex RC methods resulted in slightly higher R^2's (increase <2.6%) than the simple RC method. A comparison of estimated in-sample and out-of-sample R^2 for all RC models based on data from the simulation study is provided in Appendix Table A4.

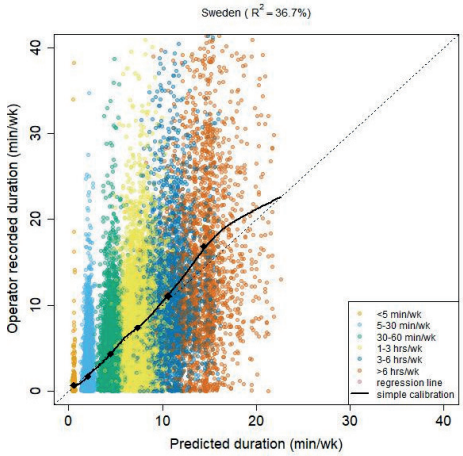
Figure 3. Observed versus predicted duration of outgoing calls by country for the direct (A) and inverse (B) RC approaches. The regression line was estimated allowing for a non-linear relation between observed and predicted values and allowing the (residual) variance in observed durations to depend on predicted duration using penalized splines (P-splines) as implemented in the gamlss software[29]. Note the different horizontal and vertical scales for the Netherlands.

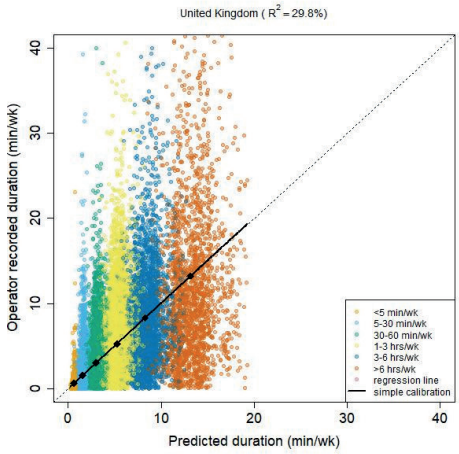


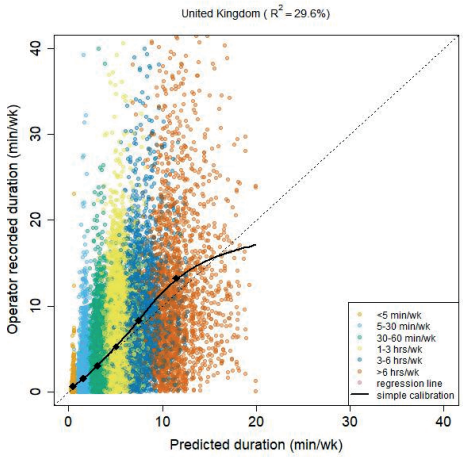

Netherlands (R2 = 32.7%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration


Netherlands (R2 = 32.4%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration


Sweden (R2 = 37.1%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration


Sweden (R2 = 36.7%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration


United Kingdom (R2 = 29.8%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration


United Kingdom (R2 = 29.6%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration

PERFORMANCE OF RC MODELS ON THE SIMULATED DATA

Simulation results are presented in Table 3. This table shows the (range of) estimated coefficients across simulations, as well as the squared bias and variance that contribute to total MSE. Using any of the RC approaches resulted in MSE's that were approximately 50% lower than those from CC analyses. Direct RC resulted in lowest MSE in most countries except Denmark, where the GAMLSS-based RC outperformed other RC approaches. MI was clearly inferior to all RC approaches. Both simple RC and inverse RC performed relatively well on data from Finland, Sweden and the UK. MSE's were lower when a larger training set was used (50% of the sample; Appendix Table A2A), especially for the inverse RC approach that performed well in almost all countries in set of simulations. Bias was a minor contributing factor to total MSE when the regression coefficient for RECORD used in the simulations was relatively small (0.001), but was more important when a larger coefficient was used (0.005; Appendix Table A2B). Point estimates obtained using GAMLSS-based RC tended to suffer from relatively large downward bias.

Table 3. Point estimates, squared bias, variance and MSE for approaches based on the results from 1,000 simulations (100 simulations for the Inverse RC approach). Simulations used data from 25% of the participants with complete information on both self-reported and operator-recorded data as the training set and the remainder as test set. The outcome was simulated using a slope coefficient (β) of 0.001 (i.e. assuming an Odds Ratio of exp(0.1)=1.11 for each additional 100 minutes call-time per week) and with a balanced ratio of cases:non-cases. Bootstrapping was used to estimate 95%CIs for each statistic. The training set was used to fit the health model for the CC analysis, to fit the first stage (exposure) models for the RC approaches, and to fit the multiple imputation model. All second stage (health) models for the RC approaches were fitted to the test data only and results were precision-weighted with those from the CC approach before further analyses. The (non-full data) model that achieved lowest MSE and all models for which the estimated MSE fell within the 95%CI for that lowest MSE are highlighted in grey.

Country	Model	b (*1000) [95%CI]	Percentiles$ (2.5%,97.5%)	Bias² [95%CI]	Variance [95%CI]	MSE# [95%CI]
Denmark	Full data	1.02	(0.34, 1.72)	0.00	0.18	0.18
		[0.99, 1.05]		[0.00, 0.00]	[0.17, 0.20]	[0.17, 0.20]
	CC	1.02	(-0.35, 2.46)	0.00	0.75	0.75
		[0.97, 1.08]		[0.00, 0.00]	[0.69, 0.82]	[0.69, 0.82]
	Simple RC	0.98	(0.05, 1.93)	0.00	0.34	0.34
		[0.94, 1.01]		[0.00, 0.00]	[0.31, 0.37]	[0.31, 0.37]
	GAMLSS RC	0.85	(0.05, 1.70)	0.02	0.26	0.28
		[0.82, 0.88]		[0.01, 0.03]	[0.24, 0.29]	[0.26, 0.31]
	Direct RC	0.96	(0.09, 1.91)	0.00	0.31	0.31
		[0.92, 0.99]		[0.00, 0.01]	[0.29, 0.34]	[0.29, 0.34]
	Inverse RC	0.90	(-0.16, 1.77)	0.01	0.35	0.36
		[0.77, 1.01]		[0.00, 0.05]	[0.28, 0.46]	[0.29, 0.48]
	MI	0.95	(-0.14, 2.10)	0.00	0.48	0.48
		[0.91, 1.00]		[0.00, 0.01]	[0.44, 0.52]	[0.44, 0.53]
Finland	Full data	1.00	(0.55, 1.45)	0.00	0.08	0.08
		[0.99, 1.02]		[0.00, 0.00]	[0.07, 0.08]	[0.07, 0.08]
	CC	1.02	(0.13, 1.95)	0.00	0.31	0.31

Country	Model	b (*1000) [95%CI]	Percentiles$ (2.5%,97.5%)	Bias² [95%CI]	Variance [95%CI]	MSE# [95%CI]
		[0.98, 1.05]		[0.00, 0.00]	[0.28, 0.34]	[0.28, 0.34]
	Simple RC	1.01	(0.33, 1.70)	0.00	0.18	0.18
		[0.98, 1.04]		[0.00, 0.00]	[0.16, 0.19]	[0.16, 0.19]
	GAMLSS RC	0.79	(0.24, 1.37)	0.05	0.11	0.16
		[0.76, 0.81]		[0.04, 0.06]	[0.11, 0.12]	[0.15, 0.17]
	Direct RC	1.01	(0.34, 1.70)	0.00	0.17	0.17
		[0.98, 1.03]		[0.00, 0.00]	[0.16, 0.19]	[0.16, 0.19]
	Inverse RC	1.13	(0.40, 1.94)	0.02	0.25	0.26
		[1.02, 1.22]		[0.00, 0.05]	[0.19, 0.32]	[0.20, 0.35]
	MI	1.05	(0.16, 1.94)	0.00	0.29	0.30
		[1.01, 1.08]		[0.00, 0.01]	[0.27, 0.32]	[0.27, 0.33]
Netherlands	Full data	1.06	(-1.48, 3.83)	0.00	2.63	2.63
		[0.96, 1.16]		[0.00, 0.03]	[2.44, 2.87]	[2.43, 2.88]
	CC	1.19	(-4.02, 6.49)	0.04	10.58	10.61
		[0.99, 1.39]		[0.00, 0.15]	[9.59, 11.56]	[9.62, 11.62]
	Simple RC	1.06	(-2.70, 5.05)	0.00	5.62	5.62
		[0.91, 1.21]		[0.00, 0.03]	[5.13, 6.20]	[5.14, 6.21]
	GAMLSS RC	0.97	(-2.54, 4.64)	0.00	4.83	4.83
		[0.85, 1.12]		[0.00, 0.01]	[4.43, 5.27]	[4.43, 5.27]
	Direct RC	0.81	(-2.13, 3.93)	0.03	3.58	3.62
		[0.70, 0.94]		[0.00, 0.09]	[3.26, 3.91]	[3.28, 3.95]
	Inverse RC	1.19	(-3.24, 4.87)	0.03	5.16	5.19
		[0.77, 1.62]		[0.00, 0.37]	[3.75, 7.89]	[3.72, 7.59]
	MI	1.17	(-3.96, 5.89)	0.03	8.69	8.73
		[0.99, 1.36]		[0.00, 0.13]	[8.02, 9.55]	[8.06, 9.59]
Sweden	Full data	1.00	(0.78, 1.21)	0.00	0.02	0.02
		[1.00, 1.01]		[0.00, 0.00]	[0.02, 0.02]	[0.02, 0.02]
	CC	1.00	(0.57, 1.46)	0.00	0.07	0.07
		[0.99, 1.02]		[0.00, 0.00]	[0.06, 0.08]	[0.06, 0.08]
	Simple RC	1.00	(0.72, 1.28)	0.00	0.03	0.03
		[0.98, 1.01]		[0.00, 0.00]	[0.03, 0.03]	[0.03, 0.03]
	GAMLSS RC	0.78	(0.56, 1.02)	0.05	0.02	0.07
		[0.77, 0.79]		[0.04, 0.05]	[0.02, 0.02]	[0.06, 0.07]
	Direct RC	0.99	(0.72, 1.28)	0.00	0.03	0.03
		[0.98, 1.00]		[0.00, 0.00]	[0.03, 0.03]	[0.03, 0.03]
	Inverse RC	1.04	(0.80, 1.28)	0.00	0.03	0.03
		[1.01, 1.08]		[0.00, 0.01]	[0.02, 0.04]	[0.02, 0.04]
	MI	1.05	(0.62, 1.47)	0.00	0.07	0.07
		[1.03, 1.06]		[0.00, 0.00]	[0.06, 0.07]	[0.06, 0.08]
United Kingdom	Full data	1.00	(0.76, 1.25)	0.00	0.02	0.02
		[0.99, 1.01]		[0.00, 0.00]	[0.02, 0.02]	[0.02, 0.02]
	CC	1.00	(0.53, 1.51)	0.00	0.09	0.09
		[0.98, 1.02]		[0.00, 0.00]	[0.08, 0.10]	[0.08, 0.10]
	Simple RC	0.99	(0.64, 1.35)	0.00	0.05	0.05

Country	Model	b (*1000) [95%CI]	Percentiles$ (2.5%,97.5%)	Bias² [95%CI]	Variance [95%CI]	MSE# [95%CI]
		[0.97, 1.00]		[0.00, 0.00]	[0.04, 0.05]	[0.04, 0.05]
	GAMLSS RC	0.92	(0.60, 1.26)	0.01	0.04	0.05
		[0.91, 0.93]		[0.00, 0.01]	[0.04, 0.05]	[0.04, 0.05]
	Direct RC	0.98	(0.64, 1.34)	0.00	0.05	0.05
		[0.97, 0.99]		[0.00, 0.00]	[0.04, 0.05]	[0.04, 0.05]
	Inverse RC	1.01	(0.62, 1.38)	0.00	0.06	0.06
		[0.97, 1.06]		[0.00, 0.00]	[0.05, 0.08]	[0.05, 0.08]
	MI	1.05	(0.55, 1.56)	0.00	0.09	0.09
		[1.03, 1.07]		[0.00, 0.00]	[0.08, 0.09]	[0.08, 0.10]

$ (Empirical) quantiles of the distribution of b estimates across simulations
# MSE = mean squared error (bias² + variance)
CC, complete case; CI, confidence interval; MI, multiple imputation; RC, regression calibration; RR, relative risk.

Coverage of Wald-type 95% confidence intervals based on estimated standard errors after inverse-precision weighting is shown in (Table 4). Coverage was nominal (i.e 95%) for the simple RC and direct RC methods, tended to be slightly too low for the inverse RC method (~93%), but was much lower for the GAMLSS-based RC in some countries (e.g. only 66% for Sweden). The average ratio of estimated slope (β) over standard error (SE) was highest for the direct RC or inverse RC methods for all countries, confirming higher statistical efficiency of these methods over alternative approaches.

Table 4. Coverage of 95%CI and ratio of slope estimate (β) over its standard error (SE) as a proxy for efficiency modelling approaches based on the results from 1,000 simulations (100 simulations for the Inverse RC approach). Simulations used data from 25% of the participants with complete information on both self-reported and operator-recorded data as the training set and the remainder as test set. The outcome was simulated using a slope coefficient (b) of 0.001 (i.e. assuming an Odds Ratio of exp(0.1)=1.11 for each additional 100 minutes call-time per week) and with a balanced ratio of cases: non-cases. Bootstrapping was used to estimate 95%CIs for each statistic.

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
Denmark	Full data	95%	2.37	(0.82, 3.91)
		[94%, 96%]	[2.31, 2.43]	
	CC	95%	1.16	(-0.41, 2.68)
		[94%, 97%]	[1.10, 1.22]	
	Simple RC	96%	1.67	(0.08, 3.23)
		[94%, 97%]	[1.61, 1.72]	
	GAMLSS RC	94%	1.66	(0.10, 3.23)
		[92%, 95%]	[1.60, 1.71]	
	Direct RC	96%	1.69	(0.15, 3.24)
		[94%, 97%]	[1.63, 1.75]	
	Inverse RC	93%	1.62	(-0.31, 3.07)
		[84%, 96%]	[1.40, 1.81]	
	MI	97%	1.28	(-0.19, 2.77)
		[95%, 97%]	[1.23, 1.34]	

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
Finland	Full data	95%	3.58	(2.02, 5.06)
		[94%, 96%]	[3.52, 3.64]	
	CC	96%	1.78	(0.25, 3.25)
		[94%, 97%]	[1.72, 1.84]	
	Simple RC	96%	2.38	(0.81, 3.92)
		[94%, 97%]	[2.32, 2.44]	
	GAMLSS RC	89%	2.33	(0.75, 3.92)
		[87%, 90%]	[2.27, 2.39]	
	Direct RC	96%	2.42	(0.80, 3.95)
		[95%, 97%]	[2.36, 2.48]	
	Inverse RC	92%	2.51	(0.88, 4.23)
		[83%, 95%]	[2.30, 2.71]	
	MI	94%	1.90	(0.31, 3.58)
		[92%, 95%]	[1.84, 1.96]	
Netherlands	Full data	96%	0.64	(-0.90, 2.29)
		[95%, 97%]	[0.58, 0.70]	
	CC	97%	0.35	(-1.22, 1.86)
		[95%, 98%]	[0.29, 0.40]	
	Simple RC	95%	0.44	(-1.15, 2.11)
		[94%, 96%]	[0.38, 0.51]	
	GAMLSS RC	95%	0.44	(-1.19, 2.05)
		[93%, 96%]	[0.38, 0.51]	
	Direct RC	94%	0.44	(-1.20, 2.12)
		[92%, 95%]	[0.38, 0.51]	
	Inverse RC	94%	0.53	(-1.43, 2.06)
		[85%, 97%]	[0.35, 0.71]	
	MI	97%	0.37	(-1.20, 1.85)
		[96%, 98%]	[0.31, 0.42]	
Sweden	Full data	95%	7.70	(6.04, 9.22)
		[94%, 96%]	[7.64, 7.76]	
	CC	95%	3.83	(2.22, 5.46)
		[93%, 96%]	[3.78, 3.89]	
	Simple RC	96%	5.57	(4.02, 7.14)
		[95%, 97%]	[5.50, 5.62]	
	GAMLSS RC	66%	5.50	(3.97, 7.04)
		[63%, 69%]	[5.45, 5.57]	
	Direct RC	96%	5.62	(4.08, 7.20)
		[95%, 97%]	[5.56, 5.68]	
	Inverse RC	96%	5.53	(4.30, 6.85)
		[89%, 98%]	[5.38, 5.69]	
	MI	94%	4.04	(2.26, 6.25)
		[92%, 95%]	[3.96, 4.12]	

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
United Kingdom	Full data	95%	6.70	(5.21, 8.24)
		[94%, 96%]	[6.65, 6.76]	
	CC	95%	3.31	(1.80, 4.96)
		[93%, 96%]	[3.25, 3.37]	
	Simple RC	94%	4.60	(3.02, 6.24)
		[93%, 95%]	[4.54, 4.66]	
	GAMLSS RC	93%	4.63	(3.09, 6.27)
		[91%, 95%]	[4.57, 4.69]	
	Direct RC	95%	4.64	(3.08, 6.26)
		[93%, 96%]	[4.58, 4.70]	
	Inverse RC	93%	4.48	(2.74, 6.12)
		[85%, 96%]	[4.28, 4.68]	
	MI	95%	3.47	(1.80, 5.30)
		[94%, 96%]	[3.41, 3.54]	

$ (Empirical) quantiles of the distribution of b/SE ratios across simulations
CC, complete case; CI, confidence interval; MI, multiple imputation; RC, regression calibration; RR, relative risk

DISCUSSION

We evaluated the performance of different approaches that combine gold-standard operator-recorded mobile phone use and its more widely available, but error-prone proxy, self-reported use, in a simulation study using actual COSMOS data with a simulated binary health outcome.

The results indicate that bias was a minor contributing factor to MSE of estimated slopes, when health outcomes were simulated under the assumption that the odds of disease increased by 2% (Odds Ratio=1.02) for each additional 20 minutes of mobile phone call-time per week. Bias became more noticeable, when outcomes were simulated assuming a stronger exposure-outcome relation, i.e. assuming an Odds Ratio of 1.11 per additional 20 minutes per week. The direction of bias (upwards or downwards) varied by country and RC approach. RC is expected to produce unbiased point estimates only for linear models[32], but our simulations confirm that bias is still limited for logistic regression models when the effect of (error-prone) exposure is not very strong or measurement error is relatively low by comparison[22,33,34], even with heteroscedastic error.

Total MSE of estimated slope coefficients were lowest for the full-data analyses and highest for the CC and MI methods, with relatively minor differences between different RC approaches, albeit with some exceptions. Among RC approaches, bias tended to be largest for the GAMLSS-based method, especially when used on the larger Swedish and UK datasets. The simple RC approach, using only self-reported mobile phone use information, performed only slightly worse than approaches relying on more complex models. This may be explained by the fact that covariates did not explain much

of the variation in RECORD once REPORT was considered, as suggested by the modest increase in average R^2 for more complex models when compared to simple RC. An advantage of using simple RC is that no additional covariates need to be available, while any gain from using more covariates would be (partly) offset by the need to consider these covariates as potential confounders in the health-outcome model[35,36]. We did not investigate this tradeoff in this study, where health outcomes were simulated using only RECORD, but this could be of practical relevance when some of the model predictors (e.g. sex and age) are also known risk factors.

RC models were fitted using two-stage regression, and standard errors from the second stage model may not fully account for the uncertainty in exposure estimates from the first stage models. Standard errors could be estimated through (asymptotic) approximation or bootstrapping[22,32], but for practical reasons we used model-based standard errors for the simulations in this paper. Coverage of 95% CIs calculated was only slightly lower than nominal (i.e. 95%) for most approaches, except for the GAMLSS-based RC, where bias in point estimates was substantial. We did investigate bootstrapped standard errors for RC models fitted to the first 100 simulations, but found these to be similar to model-based standard errors (not shown).

Simulated health outcomes were included in the MI models, but we did not consider fitting a full measurement error model that incorporates the health outcome regression[37]. Feedback between the health outcome model and the exposure and measurement error models could result in a more efficient analysis of the exposure-outcome relation, but could also result in bias amplification from model misspecification[38]. A full measurement error model would also require re-fitting the model separately for each health outcome or including all health outcomes in a single model, which is unattractive for a multi-center cohort study.

Other authors have suggested methods to either account for measurement error in self-reported mobile phone data or to assess the potential impact on study findings. Redmayne et al.[16] used a Bayesian forecasting method to correct for measurement error in the self-reported number of text messages. Their method shows similarities to the inverse RC approach applied here, but their model did not include any covariates and self-reported use was on a continuous scale. Vergnaud et al.[39] imputed missing information on Terrestrial Truncated Radio (TETRA) use among police forces using machine learning techniques. They did not include health outcomes in their imputation model and primarily reported results based on personal exposure estimates averaged across 10,000 imputations, which made it similar to our GAMLSS-based RC method. They found, as we did, that this model tended to under-estimate use for higher exposures.

Operator-recorded data was available for a quarter of participants (n=72,189, 25.6%), and was collected prior to disease occurrence. This addresses a major limitation of earlier studies, although

missing provider information could be associated with some of the health outcomes of interest. Operator data can also be problematic as they are collected by providers for billing or security rather than scientific purposes. In some countries, calls made between participants using the same provider were registered as outgoing calls only, so some incoming calls may be missing. We therefore based our simulation study on outgoing call duration, even though we may choose to use calibrated call duration estimates based on combined incoming and outgoing call duration data for the actual Cosmos analyses. There were large differences between countries regarding the amount of available operator-recorded data. Only a single provider supplied information in the Netherlands at baseline, where subjects were sampled from an existing cohort among nurses and the general population, while in the other countries sampling was (mainly) through mobile phone operator records. Collection of operator-recorded data within COSMOS will likely become more difficult over time due to attrition and participants changing network operators. Self-reported information on mobile phone use will therefore remain important[10], but may require the use of structural measurement error models to account for time-trends during follow-up, ever-changing patterns in mobile phone use, and newer network technologies. Notably, there are limitations inherent in capturing evolving mobile phone technology changes over time that may an impact on estimating exposure-response relations, which cannot be addressed with our approach. A major issue is how well mobile phone use predicts the exposure of interest, namely radiofrequency electromagnetic fields. While past validation studies have been carried out for the 2nd mobile phone generation (GSM in Europe), showing fair agreement between amount of use and cumulated emission from the handset[4], less data are available on the predictive power of mobile phones of the 3rd and the 4th generations (UMTS, LTE) that are already used by COSMOS participants.

Conclusion

This study addressed an important concern within environmental epidemiology: how to leverage self-reported exposure estimates that are often available but error-prone, with more objective measurements that may be obtained in only a subset of subjects. Our simulation study indicated RC approaches can be used to improve estimation of exposure-outcome relations between mobile phone use and health outcomes within COSMOS. The prospective design and improved exposure assessment within COSMOS are expected to allow more precise conclusions about possible health effects of contemporary exposure from mobile phones.

Funding

The Swedish part of COSMOS has been funded by the Swedish Research Council (50096102), AFA Insurance (T-26:04), the Swedish Research Council for Health, Working Life and Welfare (2010-0082, 2014-0889), and VINNOVA (P31735-1). VINNOVA received funds for this purpose from TeliaSonera AB, Ericsson AB and Telenor Sverige AB, to cover part of the data collection (ended 2012). The provision of funds to the COSMOS study investigators via VINNOVA was governed by agreements that guarantees COSMOS' complete scientific independence. TeliaSonera, Telenor, 3, and Tele2 made it possible for their subscribers to participate with traffic data. The UK part of COSMOS was funded for an initial 5 year period by the MTHR (Mobile Telecommunications and Health Research), an independent programme of research into mobile phones and health that was jointly supported by the UK Department of Health and the mobile telecommunications industry (project reference number 091/0006) and, subsequently, funded by the UK Department of Health via its Policy Research Programme (project reference number PR-ST-0713-00003). The UK research was also part funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Health Impact of Environmental Hazards a partnership between King's College London, Imperial College London and Public Health England (PHE) (HPRU-2012-10141). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR, the UK Department of Health or Public Health England. The Finnish cohort was supported by a grant from the National Technology Agency (TEKES), with contributions to the research programme from Nokia, TeliaSonera and Elisa; Pirkanmaa Hospital District competitive research funding (grant no. VTR 9T003); Yrjö Jahnsson Foundation (grant no. 5692); and an unrestricted grant from Mobile Manufacturers' Forum (with Pirkanmaa Hospital District as a firewall) with a contract guaranteeing the complete scientific independence of the researchers to analyse, interpret and report the results with no influence for the funding sources. The Dutch part of COSMOS was funded by the Netherlands Organisation for Health Research and Development (ZonMw) within the Electromagnetic Fields and Health Research programme (grant numbers 85200001, 85500003 and 85800001). The Danish part of COSMOS was funded by the Danish Strategic Research Council (grants 2103-05-0006/2064-04-0010). The French part of COSMOS is funded by the French Agency for Food, Environmental and Occupational Health & Safety (ANSES), project reference numbers 2013-CRD-17, 2015-CRD-30, 2018-CRD-03, 2020-CRD-RF20-01 and the International Agency for Research on Cancer.

Acknowledgements

We want to thank all the participants who have joined the COSMOS cohort study (http://www.thecosmosproject.org/). We thank mobile phone network operators in Denmark, Finland, France, the Netherlands, Sweden and the UK for allowing invitation of their subscribers and/or provision of

operator traffic data. PE is Director of the MRC-PHE Centre for Environment and Health supported by the Medical Research Council and Public Health England (grant number MRC G0801056), and the UK MEDical BIOinformatics partnership (UK MED-BIO) supported by the Medical Research Council(MR/L01632X/1). PE acknowledges funding from the NIHR Imperial Biomedical Research Centre, the NIHR Health Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-10141) and the UK Dementia Research Institute supported by UK DRI Ltd which is funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK. PE is associate director of Health Data Research UK-London which receives funding from a consortium led by the UK Medical Research Council. MR, LP, RV and HK were supported by The Netherlands Organization for Health Research (ZonMW) within the programme Electromagnetic Fields and Health Research, under grant numbers 85200001, 85500003 and 85800001. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer /World Health Organization.

SUPPLEMENTARY METHODS CHAPTER 3

We start with describing the health outcome model used for the *complete case* (CC) and second stage analyses below, followed by a description of the different *regression calibration* (RC) approaches used to build the first stage models and the *multiple imputation* (MI) procedure.

Health outcome model

We regard the operator-recorded mobile phone use data as the gold standard for *"true"* mobile phone use, and the goal of our analyses is therefore to estimate the slope coefficient (β) in the following health outcome model:

$$g(E(Y)) = a + \beta.RECORD + \gamma.C \qquad \text{(eq. I)}$$

where g is the logistic link function in case of a binary outcome Y, RECORD is operator-recorded mobile phone use, and C is a matrix of known or suspected confounders. Non-linear effects of either exposure or confounders can be accommodated using basis-expansions methods (f.i. using a cubic spline basis for RECORD or age). We did not include any scenarios that involve confounding or non-linear effects in the present simulations, but it should be noted that predictors used in the first stage (exposure) models should be considered as potential confounders by design[1].

Complete case analysis

For the *complete case* analysis, we fitted the health outcome model (eq. I) using data from participants in the training dataset only. Unless the number of participants without operator-recorded phone use information is low, this option is likely to be inefficient, resulting in high standard errors for the regression coefficient (β). When operator-recorded information is not missing completely at random (MCAR), complete case analysis could also result in biased estimation of this regression coefficient. For the simulations in this paper, random sampling was used to assign participants to either training or test datasets, and therefore the MCAR assumption holds and no bias is expected.

Regression calibration approaches

We consider self-reported mobile phone use (REPORT) as an (error-prone) proxy for operator-recorded use (RECORD) and consider four different implementations of *regression calibration*[2] (RC) to correct for measurement error. This involves fitting the so-called *"calibrated"* health outcome model:

$$g(E(Y)) = a^* + \beta^*.E(RECORD|REPORT,Z) \quad + \gamma^*.C \qquad \text{(eq. II)}$$

where E(RECORD|REPORT,Z) is the expected value of *"true"* operator-recorded mobile phone use, conditional on REPORT and other covariates (Z). This model may be used to adjust slope estimates

from a health outcome model fitted to the error-prone exposure proxy, but adjusted slope estimates can also be obtained through two-stage regression, as we do here. In that case, the *"calibrated"* health outcome model (eq. II) is the second-stage model, while E(RECORD|REPORT,Z) is estimated from the regression of RECORD on REPORT and other covariates in the first-stage model using the training sample:

$$E(RECORD|REPORT,Z) = \phi + h.REPORT + \lambda.Z \qquad \text{(eq. III)}$$

Our first approach, simple regression calibration (simple RC), does not use any covariates and estimates E(RECORD|REPORT) by simply averaging operator-recorded mobile phone use (RECORD) across different levels of the REPORT categories.

For the other variants we chose Z to include sex, age, educational achievement, employment status, and marital status of the study participants. As for the health outcome model, non-linear effects of continuous covariates can be accommodated using basis expansion methods, and we use a three degrees of freedom cubic spline basis for age here as well. The distribution of RECORD is strongly skewed to the right, even withing REPORT categories, and we therefore chose to estimate the (conditional) mean of RECORD using models for log-transformed RECORD as follows:

$$E(RECORD|REPORT,Z) = \exp(E(\log(RECORD)|REPORT,Z) + Var(\log(RECORD)|REPORT,Z)/2) \qquad \text{(eq. IV)}$$

with

$$E(\log(RECORD)|REPORT,Z) = \phi_{log} + \eta_{log}.REPORT + \lambda_{log}.Z \qquad \text{(eq. V)}$$

$$Var(\log(RECORD)|REPORT,Z) = \exp(\phi_{var} + \eta_{var}.REPORT + \lambda_{var}.Z) \qquad \text{(eq. VI)}$$

Our second approach, which we call GAMLSS regression calibration (GAMLSS RC), uses software to fit generalized additive models for location, shape, and scale (GAMLSS) to estimate equations V and VI, and then uses equation IV to estimate E(RECORD|REPORT,Z).

A challenge in estimating equation IV using GAMLSS models is that a good estimate of Var(log(RECORD)|REPORT,Z) is critically important but that the (conditional) distribution of log(RECORD) is still noticably skewed across REPORT categories and covariates (Z). As an alternative we therefore used a non-parametric Bayesian approach to approximate the residual distribution by fitting a (truncated) Dirichlet process mixture of (log)normal distributions[3] to the residuals within each REPORT category and averaging across draws from that distribution. We used uninformative priors for the regression parameters and truncated the Dirichlet process to use no more than 5 mixture components per REPORT category to avoid excessive running times for the simulations, but allowed for more components (n=10) when fitting the model to actual cohort data. The models were fitted using

the Bayesian modelling software JAGS[4]. We call this (third) approach "direct" regression calibration (direct RC) to distinguish it from the next.

The GAMLSS-based RC and direct RC approaches outlined above ignore the causal relation between RECORD and REPORT, treating REPORT as a categorical predictor of RECORD instead (without introducing any constraints to guarantee monotonicity), so we considered ways to improve the model by imposing additional structure on the regression estimates, i.e. by explicitly modelling how RECORD and other covariates could influence self-reported phone use. This fourth approach, which we call inverse regression calibration (inverse RC) includes a probit regression model for the (ordinal) REPORT variable with an underlying (latent) variable (U*) that is affected by both RECORD and other covariates. The observed REPORT categories result from thresholding this variable, as follows:

$$E(U^*) = \upsilon.\text{RECORD} + \omega.Z \qquad \text{(eq. VII)}$$

with $(\tau_i < U^* \leq \tau_{i+1})$ implying that REPORT = i

We allow for possible non-linear effects of RECORD and components of Z by using basis expansion methods as in the other models. To allow for missing (unobserved) RECORD variables, we estimate the distribution of RECORD semi-parametrically, as a truncated Dirichlet process mixture of normal distributions which we truncate at 5 components for the simulations (and 10 for fits to the full cohort data). To obtain estimates of RECORD conditional on categories of REPORT and other covariates from this model, we need to inverse the probit model, which can be done relatively easily within a Bayesian framework. We used flat priors for the regression parameters and fitted the models using JAGS[4].

Because health data are available for all participants that contribute to the first-stage model (i.e. have operator-recorded data), we follow the approach suggested by Spiegelman et al.[2] for studies that have an internal validation study. This approach consists of combining the slope estimate from the complete-case model fit to the internal validation data (the training set) with that from the second stage model (the "calibrated" health model) fit to the remainder of the data (the test set), using inverse-precision weighting. For the actual COSMOS analsysis, where it may be important to check for differences in exposure-response between subjects with or without operator-recorded data, one could assign the calibrated phone use data only to participants in the test set and use the operator-recorded phone use for participants in the training set to allow more direct testing of slope equality and other transportability checks.

Multiple imputation

Measurement error can be regarded as a missing data issue[5], and we therefore considered multiple imputation (MI) in addition to the complete-case and RC analyses for the simulations. MI of

operator-recorded phone use was performed by chained equations (fully conditional specification) using predictive mean matching as implemented in the R package mice[6]. Participant characteristics considered as potential predictors in the imputation model were the same as for the Indirect RC approach (i.e. self-reported mobile phone use, sex, age, educational achievement, employment status, and marital status), but additionally included the (simulated) health outcome (Y). We allowed for a potential non-linear relation between age and RECORD by using a cubic spline basis expansion for age with 3 degrees of freedom. We imputed a total of 10 different datasets and combined exposure slope coefficients from the health outcome model using Rubin's rule[7].

Supplementary references

1. Szpiro AA, Paciorek CJ. Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics* 2013;**24**:501–17. doi:10.1002/env.2233
2. Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* 1997;**65**:1179S-1186S. doi:10.1093/ajcn/65.4.1179S
3. Richardson S, Leblond L, Jaussent I, *et al.* Mixture models in measurement error problems, with reference to epidemiological studies. *J R Stat Soc Ser A Stat Soc* 2002;**165**:549–66. doi:10.1111/1467-985X.00252
4. Plummer M. A program for analysis of Bayesian graphical models using Gibbs sampling. *Proc 3rd Int Work Distrib Stat Comput* 2003;:20–22.
5. Blackwell M, Honaker J, King G. A Unified Approach to Measurement Error and Missing Data: Overview and Applications. *Sociol Methods Res* 2017;**46**:303–41. doi:10.1177/0049124115585360
6. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations. *J Stat Softw* 2011;**16**:1–67.
7. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. 1987. doi:10.1002/9780470316696

SUPPLEMENTARY TABLES AND FIGURES CHAPTER 3

Table A1. Comparison of subject characteristics for participants in the COSMOS study that either had or did not have information available on subject covariate data, self-reported mobile phone use, and operator-recorded data on mobile phone use (duration of calls), by country. P values in the last column indicate statistical significance of differences between subjects with covariate and self-reported mobile phone use data that either had or did not have operator-recorded data available (the two penultimate columns) and is based on the chi-square test for categorical variables and the Wilcoxon rank-test for age and recorded call duration.

	Denmark					
	Covariate data missing	**Self-reported & operator-recorded data missing**	**Non-users or self-reported data missing**	**Operator-recorded data missing**	**Complete data available**	**P**
Number of participants (n)	706	1859	12	20342	2993	
Age						
- No. observations	706 [100%]	-	-	-	-	
- Years (mean [sd])	50.9 [10.7]	57.3 [9.7]	53.4 [14.2]	50.7 [11.2]	46.9 [12.9]	<0.001
Sex (n [%])						
- No. observations	706 [100%]	-	-	-	-	
- Male	285 [40%]	797 [43%]	6 [50%]	10504 [52%]	1350 [45%]	<0.001
- Female	421 [60%]	1062 [57%]	6 [50%]	9838 [48%]	1643 [55%]	
Marital status (n [%])						
- No. observations	380 [54%]	-	-	-	-	
- Living together	215 [57%]	1468 [79%]	5 [42%]	15712 [77%]	2004 [67%]	<0.001
- Not living together	47 [12%]	62 [3%]	0 [0%]	1308 [6%]	321 [11%]	
- Not in a relationship	118 [31%]	329 [18%]	7 [58%]	3322 [16%]	668 [22%]	
Employment status (n [%])						
- No. observations	240 [34%]	-	-	-	-	
- Active	165 [69%]	1062 [57%]	8 [67%]	15538 [76%]	2212 [74%]	<0.01
- Inactive	75 [31%]	797 [43%]	4 [33%]	4804 [24%]	781 [26%]	
Educational level (n [%])						
- No. observations	586 [83%]	-	-	-	-	
- Elementary school	117 [20%]	235 [13%]	4 [33%]	2608 [13%]	423 [14%]	0.05
- At least secondary school	469 [80%]	1624 [87%]	8 [67%]	17734 [87%]	2570 [86%]	
Self-reported call duration (n [%])						
- No. observations	647 [92%]	-	-	-	-	
< 5 min/week	49 [8%]	-	-	1484 [7%]	132 [4%]	<0.001
5-29min/week	230 [36%]	-	-	8057 [40%]	1086 [36%]	
30-59 min/week	131 [20%]	-	-	4334 [21%]	661 [22%]	
1-3 hours/week	156 [24%]	-	-	4184 [21%]	676 [23%]	
4-6 hours/week	50 [8%]	-	-	1420 [7%]	268 [9%]	
>6 hours/week	31 [5%]	-	-	863 [4%]	170 [6%]	
Recorded call duration						
- No. observations	88 [12%]	-	-	-	-	
- Minutes/week (GM [GSD])	73.4 [3.7]	-	28.0 [4.1]	-	60.6 [3.3]	

	Finland					
	Covariate data missing	**Self-reported & operator-recorded data missing**	**Non-users or self-reported data missing**	**Operator-recorded data missing**	**Complete data available**	**P**
Number of participants (n)	260	101	449	3090	9162	
Age						
- No. observations	258 [99%]	-	-	-	-	
- Years (mean [sd])	55.3 [12.9]	53.0 [13.6]	53.6 [13.2]	48.2 [13.1]	49.3 [14.0]	<0.001
Sex (n [%])						
- No. observations	260 [100%]	-	-	-	-	
- Male	110 [42%]	40 [40%]	196 [44%]	1529 [49%]	3808 [42%]	<0.001
- Female	150 [58%]	61 [60%]	253 [56%]	1561 [51%]	5354 [58%]	
Marital status (n [%])						
- No. observations	208 [80%]	-	-	-	-	
- Living together	147 [71%]	59 [58%]	279 [62%]	2253 [73%]	6386 [70%]	<0.01
- Not living together	13 [6%]	6 [6%]	31 [7%]	268 [9%]	860 [9%]	
- Not in a relationship	48 [23%]	36 [36%]	139 [31%]	569 [18%]	1916 [21%]	
Employment status (n [%])						
- No. observations	123 [47%]	-	-	-	-	
- Active	46 [37%]	48 [48%]	187 [42%]	2056 [67%]	4721 [52%]	<0.001
- Inactive	77 [63%]	53 [52%]	262 [58%]	1034 [33%]	4441 [48%]	
Educational level (n [%])						
- No. observations	139 [53%]	-	-	-	-	
- Elementary school	80 [58%]	45 [45%]	231 [51%]	1145 [37%]	4014 [44%]	<0.001
- At least secondary school	59 [42%]	56 [55%]	218 [49%]	1945 [63%]	5148 [56%]	
Self-reported call duration (n [%])						
- No. observations	260 [100%]	-	-	-	-	
< 5 min/week	8 [3%]	-	-	19 [1%]	101 [1%]	<0.001
5-29min/week	42 [16%]	-	-	466 [15%]	1634 [18%]	
30-59 min/week	52 [20%]	-	-	645 [21%]	2175 [24%]	
1-3 hours/week	83 [32%]	-	-	1173 [38%]	3573 [39%]	
4-6 hours/week	28 [11%]	-	-	482 [16%]	1181 [13%]	
>6 hours/week	47 [18%]	-	-	305 [10%]	498 [5%]	
Recorded call duration						
- No. observations	206 [79%]	-	-	-	-	
- Minutes/week (GM [GSD])	84.2 [3.0]	-	46.5 [4.6]	-	81.5 [3.0]	

	Netherlands					
	Covariate data missing	**Self-reported & operator-recorded data missing**	**Non-users or self-reported data missing**	**Operator-recorded data missing**	**Complete data available**	**P**
Number of participants (n)	3428	28545	615	52839	3039	
Age						
- No. observations	3427 [100%]	-	-	-	-	
- Years (mean [sd])	64.6 [12.6]	54.4 [11.8]	49.4 [10.1]	47.2 [11.9]	45.0 [12.3]	<0.001
Sex (n [%])						
- No. observations	3428 [100%]	-	-	-	-	
- Male	438 [13%]	2353 [8%]	44 [7%]	6189 [12%]	280 [9%]	<0.001
- Female	2990 [87%]	26192 [92%]	571 [93%]	46650 [88%]	2759 [91%]	
Marital status (n [%])						
- No. observations	3069 [90%]	-	-	-	-	
- Living together	2026 [66%]	22707 [80%]	469 [76%]	41957 [79%]	2282 [75%]	<0.001
- Not living together	22 [1%]	437 [2%]	13 [2%]	1701 [3%]	131 [4%]	
- Not in a relationship	1021 [33%]	5401 [19%]	133 [22%]	9181 [17%]	626 [21%]	
Employment status (n [%])						
- No. observations	783 [23%]	-	-	-	-	
- Active	472 [60%]	17701 [62%]	485 [79%]	44041 [83%]	2562 [84%]	0.18
- Inactive	311 [40%]	10844 [38%]	130 [21%]	8798 [17%]	477 [16%]	
Educational level (n [%])						
- No. observations	2795 [82%]	-	-	-	-	
- Elementary school	1137 [41%]	4226 [15%]	46 [7%]	3306 [6%]	165 [5%]	0.07
- At least secondary school	1658 [59%]	24319 [85%]	569 [93%]	49533 [94%]	2874 [95%]	
Self-reported call duration (n [%])						
- No. observations	3428 [100%]	-	-	-	-	
< 5 min/week	338 [10%]	-	-	7757 [15%]	378 [12%]	<0.001
5-29 min/week	726 [21%]	-	-	23867 [45%]	1555 [51%]	
30-59 min/week	186 [5%]	-	-	9811 [19%]	642 [21%]	
1-3 hours/week	123 [4%]	-	-	7751 [15%]	359 [12%]	
4-6 hours/week	29 [1%]	-	-	2219 [4%]	75 [2%]	
>6 hours/week	2026 [59%]	-	-	1434 [3%]	30 [1%]	
Recorded call duration						
- No. observations	55 [2%]	-	-	-	-	
- Minutes/week (GM [GSD])	20.0 [3.9]	-	10.3 [3.2]	-	23.4 [2.6]	

	Sweden					
	Covariate data missing	**Self-reported & operator-recorded data missing**	**Non-users or self-reported data missing**	**Operator-recorded data missing**	**Complete data available**	**P**
Number of participants (n)	2641	4863	231	22855	24881	
Age						
- No. observations	2446 [93%]	-	-	-	-	
- Years (mean [sd])	47.5 [14.4]	50.0 [14.3]	48.2 [13.2]	44.8 [13.5]	43.3 [13.4]	<0.001
Sex (n [%])						
- No. observations	2446 [93%]	-	-	-	-	
- Male	925 [38%]	1976 [41%]	103 [45%]	11580 [51%]	11352 [46%]	<0.001
- Female	1521 [62%]	2887 [59%]	128 [55%]	11275 [49%]	13529 [54%]	
Marital status (n [%])						
- No. observations	1516 [57%]	-	-	-	-	
- Living together	913 [60%]	3187 [66%]	160 [69%]	15365 [67%]	16784 [67%]	<0.001
- Not living together	170 [11%]	243 [5%]	22 [10%]	2330 [10%]	2768 [11%]	
- Not in a relationship	433 [29%]	1433 [29%]	49 [21%]	5160 [23%]	5329 [21%]	
Employment status (n [%])						
- No. observations	440 [17%]	-	-	-	-	
- Active	224 [51%]	2956 [61%]	160 [69%]	17030 [75%]	18033 [72%]	<0.001
- Inactive	216 [49%]	1907 [39%]	71 [31%]	5825 [25%]	6848 [28%]	
Educational level (n [%])						
- No. observations	2012 [76%]	-	-	-	-	
- Elementary school	498 [25%]	1381 [28%]	42 [18%]	3751 [16%]	3331 [13%]	<0.001
- At least secondary school	1514 [75%]	3482 [72%]	189 [82%]	19104 [84%]	21550 [87%]	
Self-reported call duration (n [%])						
- No. observations	2641 [100%]	-	-	-	-	
< 5 min/week	159 [6%]	-	-	1501 [7%]	1021 [4%]	<0.001
5-29min/week	591 [22%]	-	-	6321 [28%]	7023 [28%]	
30-59 min/week	414 [16%]	-	-	4198 [18%]	4894 [20%]	
1-3 hours/week	530 [20%]	-	-	5536 [24%]	6647 [27%]	
4-6 hours/week	224 [8%]	-	-	2886 [13%]	3075 [12%]	
>6 hours/week	723 [27%]	-	-	2413 [11%]	2221 [9%]	
Recorded call duration						
- No. observations	1031 [39%]	-	-	-	-	
- Minutes/week (GM [GSD])	83.2 [3.9]	-	57.8 [5.1]	-	78.3 [3.9]	

	United Kingdom					
	Covariate data missing	**Self-reported & operator-recorded data missing**	**Non-users or self-reported data missing**	**Operator-recorded data missing**	**Complete data available**	**P**
Number of participants (n)	11908	6375	0	48220	32114	
Age						
- No. observations	11908 [100%]	-	-	-	-	
- Years (mean [sd])	43.0 [14.8]	50.1 [17.2]	-	44.5 [14.6]	44.2 [14.8]	0.05
Sex (n [%])						
- No. observations	11908 [100%]	-	-	-	-	
- Male	5697 [48%]	2831 [44%]	-	23970 [50%]	14365 [45%]	<0.001
- Female	6211 [52%]	3544 [56%]	-	24250 [50%]	17749 [55%]	
Marital status (n [%])						
- No. observations	2746 [23%]	-	-	-	-	
- Living together	1963 [71%]	4233 [66%]	-	32620 [68%]	20900 [65%]	<0.001
- Not living together	241 [9%]	422 [7%]	-	5922 [12%]	4259 [13%]	
- Not in a relationship	542 [20%]	1720 [27%]	-	9678 [20%]	6955 [22%]	
Employment status (n [%])						
- No. observations	2812 [24%]	-	-	-	-	
- Active	1727 [61%]	3108 [49%]	-	34135 [71%]	22374 [70%]	<0.001
- Inactive*	1085 [39%]	3267 [51%]	-	14085 [29%]	9740 [30%]	
Educational level (n [%])						
- No. observations	501 [4%]	-	-	-	-	
- Elementary school	268 [53%]	3909 [61%]	-	28525 [59%]	19563 [61%]	<0.001
- At least secondary school	233 [47%]	2466 [39%]	-	19695 [41%]	12551 [39%]	
Self-reported call duration (n [%])						
- No. observations	9400 [79%]	-	-	-	-	
< 5 min/week	373 [4%]	-	-	2629 [5%]	1238 [4%]	<0.001
5-29min/week	2252 [24%]	-	-	13506 [28%]	8749 [27%]	
30-59 min/week	1949 [21%]	-	-	10325 [21%]	7466 [23%]	
1-3 hours/week	2648 [28%]	-	-	12495 [26%]	9090 [28%]	
4-6 hours/week	1194 [13%]	-	-	5160 [11%]	3321 [10%]	
>6 hours/week	984 [10%]	-	-	4105 [9%]	2250 [7%]	
Recorded call duration						
- No. observations	2872 [24%]	-	-	-	-	
- Minutes/week (GM [GSD])	63.6 [3.4]	-	-	-	49.1 [3.4]	

*Inactive includes ritered, school/studying, unemployed, housewife/househusband, disabled

Table A2A. Point estimates, squared bias, variance and MSE for the different approaches based on the results from 1,000 simulations (100 simulations for the Inverse RC approach). Simulations used data from 50% of the participants with complete information on both self-reported and operator-recorded data as the training set and the remainder as test set. The outcome was simulated using a slope coefficient (b) of 0.001 (i.e. assuming an Odds Ratio of exp(0.02)=1.02 for each additional 20 minutes call-time per week) and with a balanced ratio of cases:non-cases. Bootstrapping was used to estimate 95%CIs for each statistic. The training set was used to fit the health model for the CC analysis, to fit the first stage (exposure) models for the RC approaches, and to fit the MI model. All second stage (health) models for the RC approaches were fitted to the test data only and results were precision-weighted with those from the CC approach before further analyses. The (non-full data) model that achieved lowest MSE and all models for which the estimated MSE fell within the 95%CI for that lowest MSE are highlighted in grey.

Country	Model	b (*1000) [95%CI]	Percentiles[$] (2.5%,97.5%)	Bias[2] [95%CI]	Variance [95%CI]	MSE[#] [95%CI]
Denmark	Full data	1.02	(0.33, 1.74)	0.00	0.19	0.19
		[0.99, 1.04]		[0.00, 0.00]	[0.17, 0.21]	[0.17, 0.21]
	CC	1.03	(0.06, 1.99)	0.00	0.35	0.35
		[0.99, 1.07]		[0.00, 0.00]	[0.32, 0.38]	[0.32, 0.38]
	Simple RC	1.01	(0.20, 1.82)	0.00	0.26	0.26
		[0.98, 1.04]		[0.00, 0.00]	[0.24, 0.28]	[0.24, 0.28]
	GAMLSS RC	0.94	(0.18, 1.74)	0.00	0.23	0.23
		[0.91, 0.97]		[0.00, 0.01]	[0.21, 0.24]	[0.21, 0.25]
	Direct RC	1.00	(0.21, 1.83)	0.00	0.25	0.25
		[0.97, 1.04]		[0.00, 0.00]	[0.23, 0.27]	[0.23, 0.27]
	Inverse RC	0.93	(0.18, 1.82)	0.00	0.24	0.24
		[0.84, 1.03]		[0.00, 0.03]	[0.18, 0.32]	[0.18, 0.32]
	MI	1.00	(0.17, 1.84)	0.00	0.28	0.28
		[0.96, 1.03]		[0.00, 0.00]	[0.26, 0.31]	[0.26, 0.31]
Finland	Full data	0.99	(0.54, 1.45)	0.00	0.08	0.08
		[0.97, 1.00]		[0.00, 0.00]	[0.07, 0.09]	[0.07, 0.09]
	CC	0.99	(0.36, 1.66)	0.00	0.16	0.16
		[0.97, 1.02]		[0.00, 0.00]	[0.14, 0.17]	[0.14, 0.17]
	Simple RC	0.99	(0.41, 1.59)	0.00	0.13	0.13
		[0.97, 1.01]		[0.00, 0.00]	[0.11, 0.14]	[0.12, 0.14]
	GAMLSS RC	0.87	(0.37, 1.40)	0.02	0.10	0.12
		[0.85, 0.89]		[0.01, 0.02]	[0.09, 0.11]	[0.11, 0.13]
	Direct RC	0.99	(0.44, 1.59)	0.00	0.12	0.12
		[0.97, 1.01]		[0.00, 0.00]	[0.11, 0.13]	[0.11, 0.13]
	Inverse RC	1.04	(0.49, 1.61)	0.00	0.13	0.13
		[0.97, 1.11]		[0.00, 0.01]	[0.11, 0.17]	[0.11, 0.17]
	MI	1.02	(0.41, 1.69)	0.00	0.15	0.15
		[0.99, 1.04]		[0.00, 0.00]	[0.14, 0.17]	[0.14, 0.17]
Netherlands	Full data	0.95	(-1.83, 3.44)	0.00	2.46	2.47
		[0.86, 1.05]		[0.00, 0.02]	[2.26, 2.68]	[2.26, 2.70]
	CC	0.94	(-2.59, 4.62)	0.00	5.09	5.09
		[0.80, 1.08]		[0.00, 0.03]	[4.65, 5.66]	[4.66, 5.67]
	Simple RC	0.94	(-2.30, 4.09)	0.00	3.76	3.76
		[0.82, 1.06]		[0.00, 0.03]	[3.44, 4.12]	[3.44, 4.12]
	GAMLSS RC	0.91	(-2.09, 3.99)	0.01	3.45	3.46
		[0.80, 1.03]		[0.00, 0.04]	[3.18, 3.76]	[3.18, 3.76]

Country	Model	b (*1000) [95%CI]	Percentiles[$] (2.5%,97.5%)	Bias[2] [95%CI]	Variance [95%CI]	MSE[#] [95%CI]
	Direct RC	0.89	(-2.12, 3.99)	0.01	3.37	3.38
		[0.76, 0.99]		[0.00, 0.06]	[3.09, 3.66]	[3.10, 3.67]
	Inverse RC	1.00	(-2.29, 4.15)	0.00	3.46	3.46
		[0.63, 1.39]		[0.00, 0.00]	[2.63, 4.59]	[2.59, 4.52]
	MI	0.96	(-2.45, 4.48)	0.00	4.47	4.47
		[0.83, 1.09]		[0.00, 0.01]	[4.09, 4.84]	[4.09, 4.86]
Sweden	Full data	1.00	(0.78, 1.21)	0.00	0.02	0.02
		[0.99, 1.01]		[0.00, 0.00]	[0.02, 0.02]	[0.02, 0.02]
	CC	1.00	(0.68, 1.32)	0.00	0.04	0.04
		[0.99, 1.01]		[0.00, 0.00]	[0.03, 0.04]	[0.03, 0.04]
	Simple RC	1.00	(0.74, 1.28)	0.00	0.03	0.03
		[0.99, 1.01]		[0.00, 0.00]	[0.02, 0.03]	[0.02, 0.03]
	GAMLSS RC	0.86	(0.64, 1.09)	0.02	0.02	0.04
		[0.85, 0.87]		[0.02, 0.02]	[0.02, 0.02]	[0.04, 0.04]
	Direct RC	1.00	(0.74, 1.27)	0.00	0.03	0.03
		[0.99, 1.01]		[0.00, 0.00]	[0.02, 0.03]	[0.02, 0.03]
	Inverse RC	1.03	(0.79, 1.25)	0.00	0.02	0.02
		[1.00, 1.06]		[0.00, 0.00]	[0.02, 0.03]	[0.02, 0.03]
	MI	1.04	(0.73, 1.34)	0.00	0.03	0.04
		[1.02, 1.05]		[0.00, 0.00]	[0.03, 0.04]	[0.03, 0.04]
United Kingdom	Full data	1.00	(0.75, 1.24)	0.00	0.02	0.02
		[0.99, 1.01]		[0.00, 0.00]	[0.02, 0.02]	[0.02, 0.02]
	CC	1.01	(0.66, 1.35)	0.00	0.04	0.04
		[1.00, 1.02]		[0.00, 0.00]	[0.04, 0.05]	[0.04, 0.05]
	Simple RC	1.00	(0.70, 1.30)	0.00	0.03	0.03
		[0.99, 1.01]		[0.00, 0.00]	[0.03, 0.04]	[0.03, 0.04]
	GAMLSS RC	0.96	(0.68, 1.25)	0.00	0.03	0.03
		[0.95, 0.97]		[0.00, 0.00]	[0.03, 0.03]	[0.03, 0.04]
	Direct RC	1.00	(0.70, 1.30)	0.00	0.03	0.03
		[0.99, 1.01]		[0.00, 0.00]	[0.03, 0.04]	[0.03, 0.04]
	Inverse RC	1.04	(0.75, 1.29)	0.00	0.03	0.03
		[1.00, 1.07]		[0.00, 0.00]	[0.02, 0.04]	[0.02, 0.04]
	MI	1.04	(0.70, 1.37)	0.00	0.04	0.04
		[1.03, 1.05]		[0.00, 0.00]	[0.04, 0.05]	[0.04, 0.05]

CC, complete case; CI, confidence interval; MI, multiple imputation; RC, regression calibration

Table A2B. Point estimates, squared bias, variance and MSE for the different approaches based on the results from 1,000 simulations (100 simulations for the Inverse RC approach). Simulations used data from 25% of the participants with complete information on both self-reported and operator-recorded data as the training set and the remainder as test set. The outcome was simulated using a slope coefficient (b) of 0.005 (i.e. assuming an Odds Ratio of exp(0.1)=1.11 for each additional 20 minutes call-time per week) and with a balanced ratio of cases:non-cases. Bootstrapping was used to estimate 95%CIs for each statistic. The training set was used to fit the health model for the CC analysis, to fit the first stage (exposure) models for the RC approaches, and to fit the multiple imputation model. All second stage (health) models for the RC approaches were fitted to the test data only and results were precision-weighted with those from the CC approach before further analyses. The (non-full data) model that achieved lowest MSE and all models for which the estimated MSE fell within the 95%CI for that lowest MSE are highlighted in grey.

Country	Model	b (*1000) [95%CI]	Percentiles[$] (2.5%,97.5%)	Bias[2] [95%CI]	Variance [95%CI]	MSE[#] [95%CI]
Denmark	Full data	5.02	(4.16, 5.90)	0.00	0.28	0.29
		[4.98, 5.05]		[0.00, 0.00]	[0.26, 0.31]	[0.26, 0.31]
	CC	5.05	(3.29, 6.84)	0.00	1.15	1.15
		[4.98, 5.12]		[0.00, 0.01]	[1.05, 1.25]	[1.05, 1.25]
	Simple RC	4.68	(3.61, 5.82)	0.10	0.46	0.56
		[4.64, 4.72]		[0.08, 0.13]	[0.42, 0.50]	[0.52, 0.61]
	GAMLSS RC	4.05	(2.99, 5.17)	0.91	0.43	1.34
		[4.00, 4.09]		[0.83, 0.99]	[0.40, 0.48]	[1.26, 1.43]
	Direct RC	4.60	(3.57, 5.73)	0.16	0.43	0.59
		[4.56, 4.64]		[0.13, 0.19]	[0.40, 0.47]	[0.54, 0.64]
	Inverse RC	4.40	(3.47, 5.37)	0.36	0.36	0.71
		[4.28, 4.52]		[0.23, 0.52]	[0.27, 0.48]	[0.57, 0.89]
	MI	4.69	(3.37, 6.12)	0.10	0.70	0.80
		[4.64, 4.75]		[0.06, 0.13]	[0.64, 0.76]	[0.73, 0.87]
Finland	Full data	5.00	(4.45, 5.55)	0.00	0.11	0.11
		[4.98, 5.02]		[0.00, 0.00]	[0.11, 0.13]	[0.11, 0.13]
	CC	5.04	(3.95, 6.20)	0.00	0.47	0.47
		[5.00, 5.08]		[0.00, 0.01]	[0.43, 0.51]	[0.43, 0.52]
	Simple RC	4.78	(4.00, 5.58)	0.05	0.24	0.29
		[4.75, 4.81]		[0.04, 0.06]	[0.22, 0.26]	[0.26, 0.31]
	GAMLSS RC	3.58	(2.76, 4.40)	2.01	0.26	2.26
		[3.55, 3.62]		[1.92, 2.10]	[0.24, 0.28]	[2.17, 2.36]
	Direct RC	4.76	(4.02, 5.52)	0.06	0.23	0.28
		[4.73, 4.79]		[0.04, 0.07]	[0.21, 0.25]	[0.26, 0.31]
	Inverse RC	5.05	(4.22, 5.91)	0.00	0.27	0.27
		[4.95, 5.16]		[0.00, 0.02]	[0.20, 0.40]	[0.20, 0.41]
	MI	5.09	(4.00, 6.27)	0.01	0.47	0.47
		[5.05, 5.13]		[0.00, 0.02]	[0.43, 0.51]	[0.43, 0.52]
Netherlands	Full data	5.10	(2.38, 7.97)	0.01	2.94	2.95
		[4.99, 5.20]		[0.00, 0.04]	[2.70, 3.21]	[2.70, 3.23]
	CC	5.18	(-0.42, 11.36)	0.03	12.76	12.80
		[4.99, 5.42]		[0.00, 0.17]	[11.71, 13.90]	[11.75, 14.02]
	Simple RC	5.05	(1.07, 9.07)	0.00	5.79	5.79
		[4.90, 5.19]		[0.00, 0.02]	[5.32, 6.31]	[5.32, 6.32]

Country	Model	b (*1000) [95%CI]	Percentiles$ (2.5%,97.5%)	Bias² [95%CI]	Variance [95%CI]	MSE# [95%CI]
	GAMLSS RC	4.57	(0.83, 8.29)	0.19	5.09	5.28
		[4.44, 4.70]		[0.09, 0.32]	[4.68, 5.60]	[4.86, 5.80]
	Direct RC	3.68	(0.57, 6.74)	1.74	3.75	5.50
		[3.56, 3.80]		[1.45, 2.07]	[3.45, 4.14]	[5.09, 5.96]
	Inverse RC	4.46	(0.64, 8.74)	0.30	6.40	6.69
		[3.98, 5.04]		[0.00, 1.03]	[4.99, 8.20]	[5.24, 8.44]
	MI	5.14	(0.31, 10.22)	0.02	9.30	9.32
		[4.95, 5.33]		[0.00, 0.11]	[8.53, 10.25]	[8.56, 10.31]
Sweden	Full data	5.00	(4.74, 5.28)	0.00	0.03	0.03
		[4.99, 5.01]		[0.00, 0.00]	[0.03, 0.03]	[0.03, 0.03]
	CC	5.01	(4.48, 5.54)	0.00	0.11	0.11
		[4.99, 5.03]		[0.00, 0.00]	[0.10, 0.12]	[0.10, 0.12]
	Simple RC	4.71	(4.36, 5.10)	0.08	0.05	0.13
		[4.70, 4.73]		[0.07, 0.09]	[0.05, 0.06]	[0.13, 0.14]
	GAMLSS RC	3.59	(3.17, 4.00)	1.99	0.06	2.05
		[3.57, 3.60]		[1.95, 2.03]	[0.06, 0.07]	[2.01, 2.10]
	Direct RC	4.69	(4.34, 5.08)	0.09	0.05	0.14
		[4.68, 4.71]		[0.09, 0.10]	[0.05, 0.06]	[0.14, 0.15]
	Inverse RC	5.12	(4.70, 5.58)	0.01	0.08	0.09
		[5.06, 5.17]		[0.00, 0.03]	[0.06, 0.11]	[0.07, 0.13]
	MI	5.25	(4.70, 5.84)	0.06	0.13	0.19
		[5.23, 5.27]		[0.05, 0.07]	[0.11, 0.14]	[0.17, 0.20]
United	Full data	5.00	(4.69, 5.32)	0.00	0.04	0.04
Kingdom		[4.99, 5.01]		[0.00, 0.00]	[0.03, 0.04]	[0.03, 0.04]
	CC	5.04	(4.38, 5.70)	0.00	0.16	0.16
		[5.01, 5.06]		[0.00, 0.00]	[0.14, 0.17]	[0.14, 0.17]
	Simple RC	4.69	(4.24, 5.13)	0.10	0.07	0.17
		[4.67, 4.71]		[0.09, 0.11]	[0.07, 0.08]	[0.16, 0.18]
	GAMLSS RC	4.30	(3.90, 4.71)	0.49	0.06	0.55
		[4.29, 4.32]		[0.46, 0.51]	[0.06, 0.07]	[0.52, 0.57]
	Direct RC	4.66	(4.22, 5.09)	0.12	0.07	0.19
		[4.64, 4.67]		[0.11, 0.13]	[0.06, 0.07]	[0.17, 0.20]
	Inverse RC	5.06	(4.61, 5.52)	0.00	0.08	0.09
		[5.01, 5.12]		[0.00, 0.02]	[0.07, 0.11]	[0.07, 0.12]
	MI	5.32	(4.59, 6.01)	0.10	0.18	0.28
		[5.30, 5.35]		[0.09, 0.12]	[0.17, 0.20]	[0.26, 0.31]

CC, complete case; CI, confidence interval; MI, multiple imputation; RC, regression calibration.

Table A3A. Coverage of 95%CI and ratio of slope estimate (β) over its standard error (SE) as a proxy for efficiency for the different approaches based on the results from 1,000 simulations (100 simulations for the Inverse RC approach). Simulations used data from 50% of the participants with complete information on both self-reported and operator-recorded data as the training set and the remainder as test set. The outcome was simulated using a slope coefficient (b) of 0.001 (i.e. assuming an Odds Ratio of exp(0.02)=1.02 for each additional 20 minutes call-time per week) and with a balanced ratio of cases:non-cases. Bootstrapping was used to estimate 95%CIs for each statistic.

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
Denmark	Full data	95%	2.38	(0.79, 3.94)
		[93%, 96%]	[2.31, 2.43]	
	CC	97%	1.68	(0.11, 3.13)
		[95%, 98%]	[1.62, 1.74]	
	Simple RC	96%	1.95	(0.39, 3.48)
		[95%, 97%]	[1.89, 2.01]	
	GAMLSS RC	96%	1.95	(0.40, 3.51)
		[94%, 97%]	[1.89, 2.02]	
	Direct RC	96%	1.97	(0.41, 3.49)
		[95%, 97%]	[1.90, 2.03]	
	Inverse RC	94%	1.84	(0.34, 3.56)
		[86%, 97%]	[1.65, 2.03]	
	MI	96%	1.77	(0.35, 3.17)
		[95%, 97%]	[1.71, 1.83]	
Finland	Full data	95%	3.52	(1.99, 5.05)
		[93%, 96%]	[3.47, 3.58]	
	CC	95%	2.50	(0.93, 4.07)
		[93%, 96%]	[2.44, 2.56]	
	Simple RC	95%	2.79	(1.23, 4.41)
		[93%, 96%]	[2.73, 2.85]	
	GAMLSS RC	93%	2.76	(1.18, 4.36)
		[91%, 94%]	[2.70, 2.82]	
	Direct RC	95%	2.83	(1.28, 4.43)
		[93%, 96%]	[2.77, 2.89]	
	Inverse RC	96%	2.85	(1.39, 4.28)
		[89%, 98%]	[2.66, 3.03]	
	MI	95%	2.56	(1.08, 4.15)
		[94%, 96%]	[2.50, 2.62]	
Netherlands	Full data	96%	0.58	(-1.11, 2.07)
		[95%, 97%]	[0.52, 0.63]	
	CC	96%	0.40	(-1.11, 1.95)
		[94%, 97%]	[0.34, 0.46]	
	Simple RC	96%	0.46	(-1.11, 2.01)
		[95%, 97%]	[0.40, 0.52]	
	GAMLSS RC	96%	0.47	(-1.08, 1.99)
		[94%, 97%]	[0.41, 0.53]	
	Direct RC	96%	0.46	(-1.09, 2.03)
		[94%, 97%]	[0.40, 0.52]	

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
	Inverse RC	96%	0.49	(-1.22, 1.92)
		[88%, 98%]	[0.31, 0.68]	
	MI	97%	0.42	(-1.07, 1.98)
		[96%, 98%]	[0.36, 0.48]	
Sweden	Full data	95%	7.67	(6.07, 9.18)
		[93%, 96%]	[7.61, 7.72]	
	CC	94%	5.44	(3.73, 7.06)
		[92%, 95%]	[5.38, 5.49]	
	Simple RC	95%	6.34	(4.73, 8.01)
		[93%, 96%]	[6.28, 6.39]	
	GAMLSS RC	81%	6.25	(4.65, 7.90)
		[79%, 84%]	[6.19, 6.31]	
	Direct RC	94%	6.37	(4.72, 8.02)
		[93%, 95%]	[6.31, 6.43]	
	Inverse RC	95%	6.34	(5.01, 7.84)
		[86%, 98%]	[6.16, 6.52]	
	MI	94%	5.60	(3.80, 7.54)
		[92%, 95%]	[5.53, 5.67]	
United Kingdom	Full data	95%	6.71	(5.08, 8.14)
		[93%, 96%]	[6.65, 6.77]	
	CC	95%	4.77	(3.19, 6.24)
		[93%, 96%]	[4.71, 4.83]	
	Simple RC	95%	5.43	(3.86, 6.93)
		[93%, 96%]	[5.38, 5.49]	
	GAMLSS RC	94%	5.45	(3.87, 6.97)
		[93%, 96%]	[5.39, 5.50]	
	Direct RC	95%	5.45	(3.91, 6.96)
		[93%, 96%]	[5.39, 5.51]	
	Inverse RC	98%	5.47	(4.03, 6.75)
		[92%, 99%]	[5.30, 5.64]	
	MI	95%	4.85	(3.19, 6.54)
		[93%, 96%]	[4.79, 4.93]	

CC, complete case; CI, confidence interval; MI, multiple imputation; RC, regression calibration.

Table A3B. Coverage of 95%CI and ratio of slope estimate (β) over its standard error (SE) as a proxy for efficiency for the different approaches based on the results from 1,000 simulations (100 simulations for the Inverse RC approach). Simulations used data from 25% of the participants with complete information on both self-reported and operator-recorded data as the training set and the remainder as test set. The outcome was simulated using a slope coefficient (b) of 0.005 (i.e. assuming an Odds Ratio of exp(0.1)=1.11 for each additional 20 minutes call-time per week) and with a balanced ratio of cases:non-cases. Bootstrapping was used to estimate 95%CIs for each statistic.

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
Denmark	Full data	96%	9.22	(8.06, 10.36)
		[94%, 97%]	[9.17, 9.26]	
	CC	96%	4.60	(3.32, 5.73)
		[94%, 97%]	[4.56, 4.65]	
	Simple RC	89%	7.25	(5.82, 8.64)
		[87%, 91%]	[7.19, 7.29]	
	GAMLSS RC	57%	7.18	(5.78, 8.65)
		[54%, 60%]	[7.13, 7.24]	
	Direct RC	86%	7.34	(5.94, 8.73)
		[84%, 88%]	[7.28, 7.39]	
	Inverse RC	82%	7.21	(5.98, 8.67)
		[72%, 88%]	[7.07, 7.35]	
	MI	96%	4.64	(3.15, 6.44)
		[94%, 97%]	[4.58, 4.71]	
Finland	Full data	96%	14.80	(13.49, 16.05)
		[94%, 97%]	[14.75, 14.85]	
	CC	96%	7.42	(6.10, 8.70)
		[94%, 97%]	[7.37, 7.47]	
	Simple RC	90%	10.24	(8.85, 11.60)
		[88%, 92%]	[10.19, 10.30]	
	GAMLSS RC	10%	9.96	(8.43, 11.43)
		[8%, 12%]	[9.90, 10.02]	
	Direct RC	90%	10.42	(9.01, 11.85)
		[88%, 91%]	[10.37, 10.48]	
	Inverse RC	95%	10.13	(8.73, 11.22)
		[89%, 98%]	[9.97, 10.30]	
	MI	96%	7.14	(4.92, 9.88)
		[94%, 97%]	[7.05, 7.24]	
Netherlands	Full data	95%	2.99	(1.44, 4.53)
		[94%, 96%]	[2.93, 3.05]	
	CC	95%	1.48	(-0.12, 3.04)
		[93%, 96%]	[1.43, 1.55]	
	Simple RC	95%	2.10	(0.47, 3.63)
		[94%, 96%]	[2.04, 2.16]	
	GAMLSS RC	93%	2.06	(0.39, 3.64)
		[91%, 94%]	[2.01, 2.12]	
	Direct RC	85%	1.97	(0.29, 3.53)
		[82%, 87%]	[1.91, 2.03]	

Country	Model	Coverage [95%CI]	β/SE [95%CI]	Percentiles$ (2.5%, 97.5%)
	Inverse RC	95%	1.89	(0.27, 3.60)
		[87%, 97%]	[1.70, 2.12]	
	MI	96%	1.58	(0.09, 2.99)
		[94%, 97%]	[1.52, 1.64]	
Sweden	Full data	95%	30.56	(29.43, 31.76)
		[93%, 96%]	[30.51, 30.61]	
	CC	95%	15.29	(14.06, 16.48)
		[94%, 96%]	[15.24, 15.33]	
	Simple RC	65%	23.75	(22.38, 25.23)
		[62%, 68%]	[23.69, 23.81]	
	GAMLSS RC	0%	23.42	(21.88, 24.99)
		[0%, 0%]	[23.37, 23.48]	
	Direct RC	61%	23.93	(22.50, 25.43)
		[58%, 64%]	[23.88, 23.99]	
	Inverse RC	83%	23.89	(22.42, 25.52)
		[73%, 88%]	[23.69, 24.12]	
	MI	90%	14.34	(10.18, 19.70)
		[87%, 91%]	[14.16, 14.53]	
United Kingdom	Full data	94%	26.28	(24.99, 27.53)
		[92%, 95%]	[26.23, 26.32]	
	CC	94%	13.19	(11.83, 14.46)
		[92%, 95%]	[13.14, 13.24]	
	Simple RC	72%	19.54	(18.07, 21.04)
		[70%, 75%]	[19.49, 19.60]	
	GAMLSS RC	15%	19.69	(18.21, 21.21)
		[13%, 17%]	[19.64, 19.75]	
	Direct RC	67%	19.72	(18.25, 21.25)
		[64%, 70%]	[19.67, 19.78]	
	Inverse RC	92%	19.47	(17.95, 20.92)
		[83%, 95%]	[19.29, 19.65]	
	MI	92%	11.58	(7.94, 16.23)
		[91%, 94%]	[11.41, 11.73]	

CC, complete case; CI, confidence interval; MI, multiple imputation; RC, regression calibration;

Table A4. In-sample and out-of-sample estimates of R^2 (A) and slope estimates of the linear regression of observed values on regression calibration (RC) estimates (B) for the simple, GAMLSS based, direct, and inverse RC models by country and (relative) training sample size.

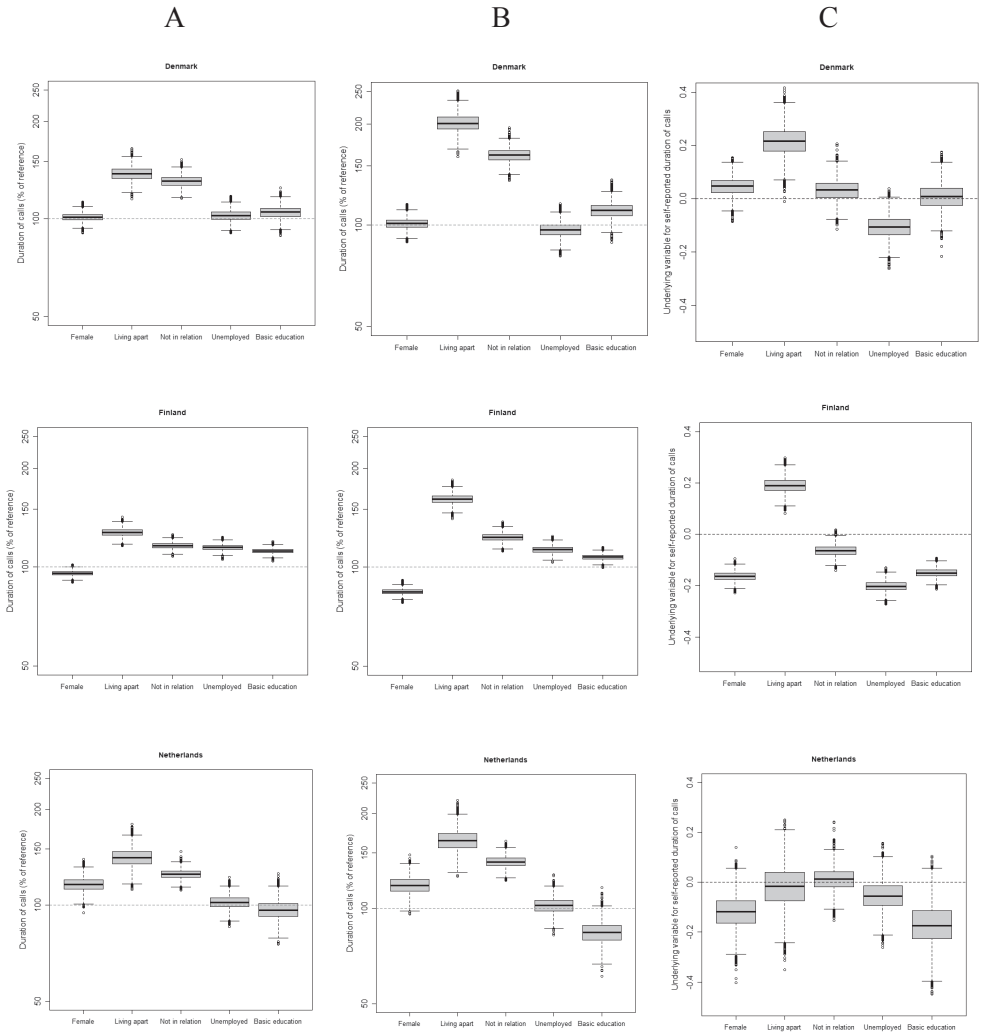
A)

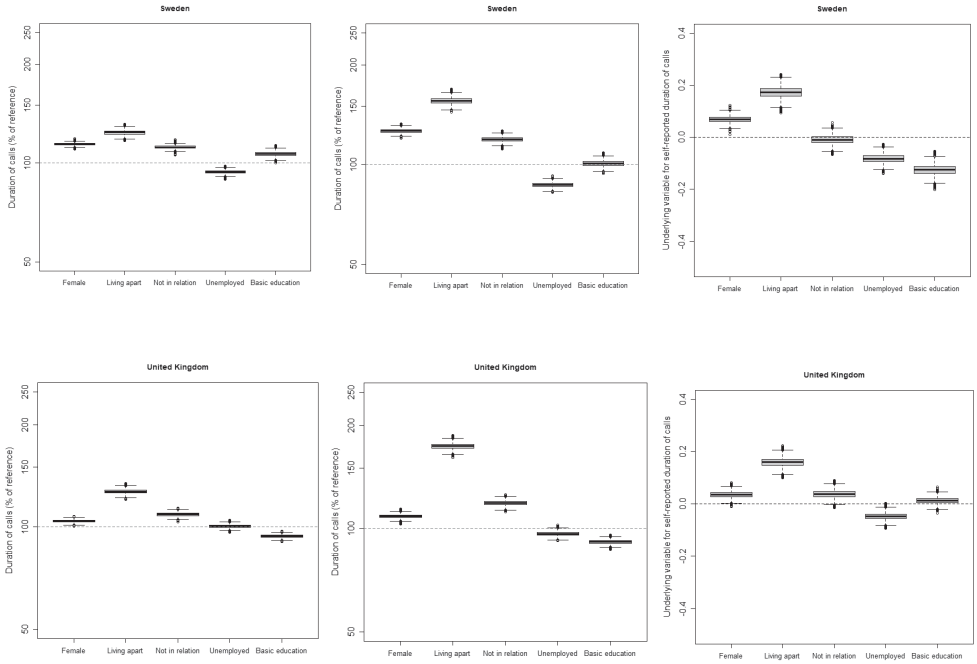
		Simple RC		GAMLSS RC		Direct RC		Inverse RC	
Country	**Training sample size**	**In-Sample [5%, 95%]**	**Out-Of-Sample [5%, 95%]**	**In-Sample [5%, 95%]**	**Out-Of-Sample [5%, 95%]**	**In-Sample [5%, 95%]**	**Out-Of-Sample [5%, 95%]**	**In-Sample [5%, 95%]**	**Out-Of-Sample [5%, 95%]**
Denmark	25%	37%	35%	38%	36%	40%	37%	39%	36%
		[31%, 43%]	[33%, 38%]	[30%, 45%]	[32%, 39%]	[34%, 46%]	[35%, 40%]	[35%, 45%]	[33%, 38%]
	50%	36%	36%	38%	37%	39%	38%	37%	37%
		[33%, 40%]	[33%, 39%]	[34%, 42%]	[33%, 41%]	[35%, 43%]	[35%, 42%]	[34%, 41%]	[34%, 42%]
Finland	25%	24%	23%	24%	24%	26%	25%	25%	24%
		[20%, 28%]	[22%, 25%]	[20%, 29%]	[21%, 26%]	[22%, 31%]	[24%, 27%]	[21%, 29%]	[22%, 26%]
	50%	24%	24%	25%	24%	26%	25%	25%	24%
		[22%, 26%]	[21%, 26%]	[22%, 27%]	[22%, 27%]	[24%, 29%]	[23%, 28%]	[22%, 28%]	[21%, 27%]
Netherlands	25%	33%	30%	33%	29%	30%	27%	32%	29%
		[27%, 38%]	[26%, 32%]	[25%, 41%]	[23%, 33%]	[23%, 37%]	[21%, 32%]	[20%, 38%]	[21%, 33%]
	50%	32%	30%	33%	31%	32%	30%	32%	30%
		[29%, 35%]	[27%, 34%]	[29%, 37%]	[27%, 35%]	[29%, 36%]	[25%, 35%]	[25%, 36%]	[23%, 35%]
Sweden	25%	36%	36%	36%	36%	37%	37%	37%	37%
		[34%, 38%]	[35%, 36%]	[34%, 39%]	[35%, 37%]	[35%, 39%]	[36%, 38%]	[35%, 39%]	[36%, 37%]
	%	36%	36%	36%	36%	37%	37%	37%	37%
		[34%, 37%]	[34%, 37%]	[35%, 38%]	[35%, 38%]	[36%, 38%]	[36%, 38%]	[36%, 38%]	[35%, 38%]
United Kingdom	25%	29%	29%	30%	30%	30%	30%	30%	30%
		[27%, 31%]	[28%, 30%]	[28%, 33%]	[29%, 31%]	[28%, 33%]	[29%, 31%]	[28%, 32%]	[29%, 31%]
	50%	29%	29%	30%	30%	30%	30%	30%	30%
		[28%, 30%]	[28%, 30%]	[29%, 32%]	[28%, 31%]	[28%, 31%]	[28%, 31%]	[28%, 31%]	[29%, 31%]

B)

Country	Training model size	Simple RC		GAMLSS RC		Direct RC		Inverse RC	
		In-Sample [5%, 95%]	Out-Of-Sample [5%, 95%]	In-Sample [5%, 95%]	Out-Of-Sample [5%, 95%]	In-Sample [5%, 95%]	Out-Of-Sample [5%, 95%]	In-Sample [5%, 95%]	Out-Of-Sample [5%, 95%]
Denmark	25%	1.00	0.99	0.83	0.80	0.99	0.96	0.92	0.90
		[1.00, 1.00]	[0.84, 1.14]	[0.66, 0.97]	[0.60, 0.98]	[0.93, 1.03]	[0.81, 1.11]	[0.81, 1.03]	[0.74, 1.09]
	50%	1.00	1.00	0.83	0.83	0.99	0.98	0.96	0.96
		[1.00, 1.00]	[0.87, 1.13]	[0.74, 0.93]	[0.69, 0.96]	[0.97, 1.02]	[0.85, 1.12]	[0.90, 1.01]	[0.84, 1.05]
Finland	25%	1.00	1.00	0.68	0.67	1.00	1.00	1.16	1.15
		[1.00, 1.00]	[0.87, 1.13]	[0.52, 0.82]	[0.51, 0.82]	[0.97, 1.05]	[0.89, 1.11]	[1.06, 1.26]	[1.00, 1.29]
	50%	1.00	1.00	0.69	0.69	1.00	1.00	1.19	1.19
		[1.00, 1.00]	[0.88, 1.12]	[0.61, 0.78]	[0.58, 0.80]	[0.98, 1.03]	[0.90, 1.10]	[1.13, 1.26]	[1.04, 1.31]
Netherlands	25%	1.00	0.97	0.88	0.83	0.64	0.61	0.98	0.90
		[1.00, 1.00]	[0.76, 1.15]	[0.59, 1.02]	[0.54, 1.04]	[0.54, 0.75]	[0.41, 0.82]	[0.63, 1.12]	[0.62, 1.12]
	50%	1.00	0.98	0.88	0.86	0.84	0.80	1.00	0.96
		[1.00, 1.00]	[0.82, 1.16]	[0.76, 0.98]	[0.67, 1.03]	[0.77, 0.90]	[0.61, 0.99]	[0.77, 1.10]	[0.70, 1.16]
Sweden	25%	1.00	1.00	0.71	0.70	1.00	0.99	1.14	1.13
		[1.00, 1.00]	[0.95, 1.05]	[0.63, 0.78]	[0.62, 0.78]	[0.99, 1.01]	[0.95, 1.04]	[1.11, 1.16]	[1.08, 1.19]
	50%	1.00	1.00	0.71	0.71	1.00	1.00	1.13	1.13
		[1.00, 1.00]	[0.96, 1.04]	[0.67, 0.75]	[0.66, 0.75]	[0.99, 1.00]	[0.95, 1.04]	[1.12, 1.15]	[1.08, 1.17]
United Kingdom	25%	1.00	1.00	0.88	0.88	0.99	0.98	1.15	1.14
		[1.00, 1.00]	[0.93, 1.06]	[0.85, 0.92]	[0.82, 0.94]	[0.97, 1.00]	[0.92, 1.04]	[1.12, 1.19]	[1.09, 1.21]
	50%	1.00	1.00	0.89	0.88	0.99	0.98	1.13	1.13
		[1.00, 1.00]	[0.94, 1.05]	[0.87, 0.91]	[0.83, 0.93]	[0.98, 1.00]	[0.93, 1.03]	[1.10, 1.15]	[1.07, 1.19]

Figure A1. Parameter estimates for the categorical predictors of the exposure models for the direct (A) and inverse (B) regression calibration (RC) approaches and for the structural measurement error model of the inverse RC approach (C) by country.
Categorical predictors: sex (female versus male), marital Status (living apart and not in a relation versus living together); employment status (inactive versus active), and educational level (elementary school only versus secondary school or higher).




Sweden
Duration of calls (% of reference)
Underlying variable for self-reported duration of calls
United Kingdom
Female
Living apart
Not in relation
Unemployed
Basic education

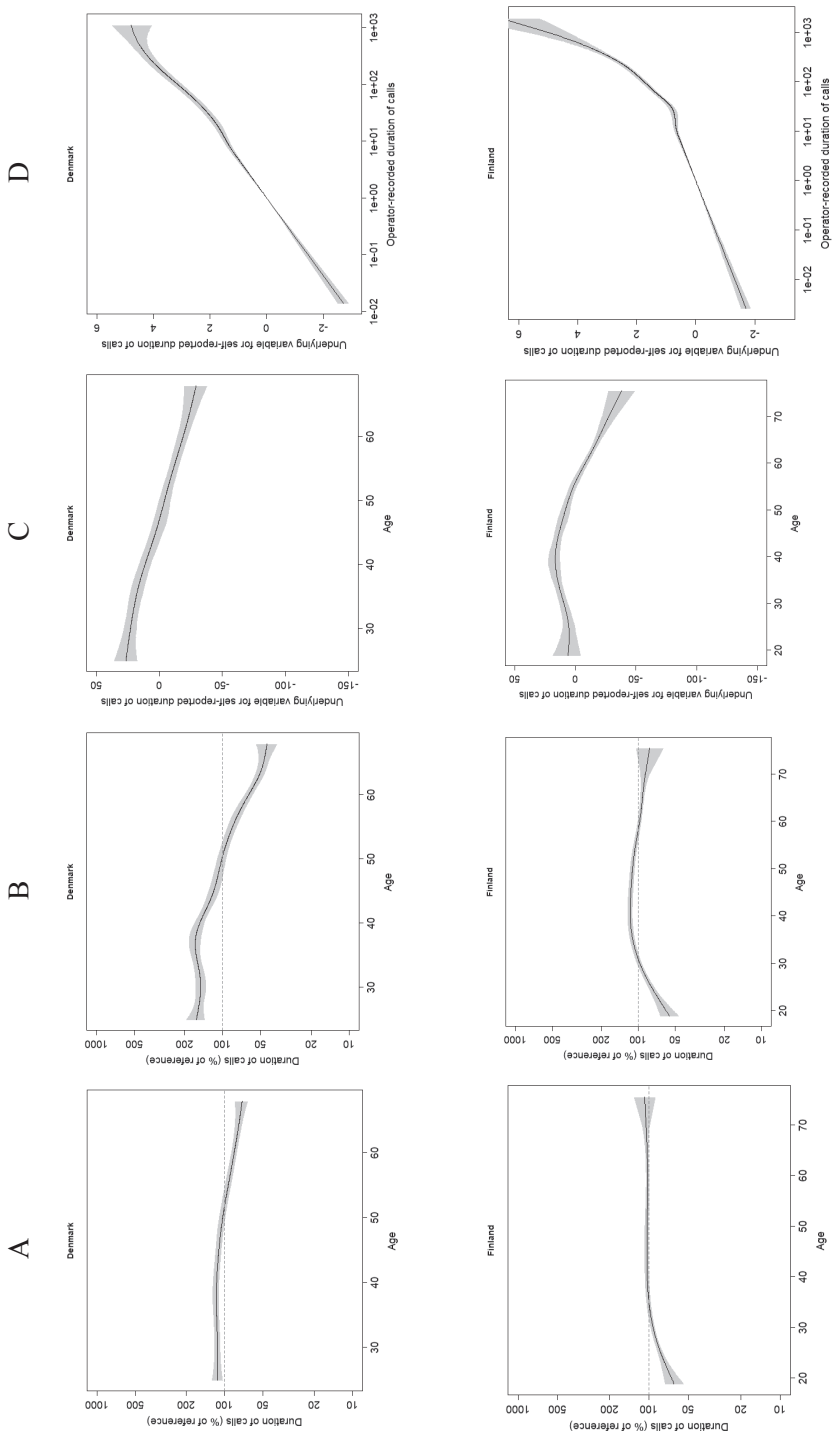


Figure A2. Effect estimates for age in the exposure models of the direct (A) and inverse (B) regression calibration (RC) approaches, and for age (C) and RECORD (D) in the structural measurement error model of the inverse RC approach by country.

3

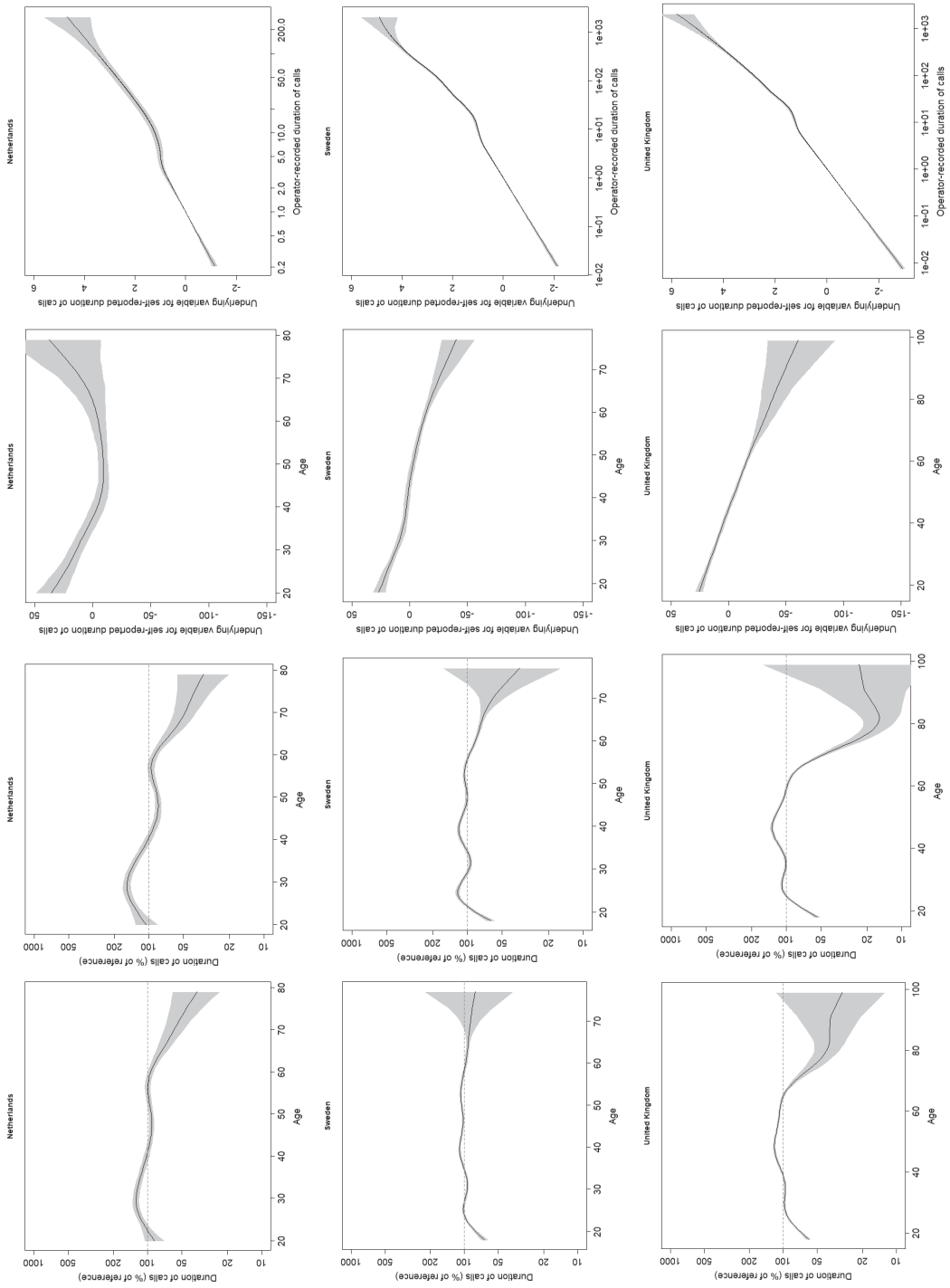
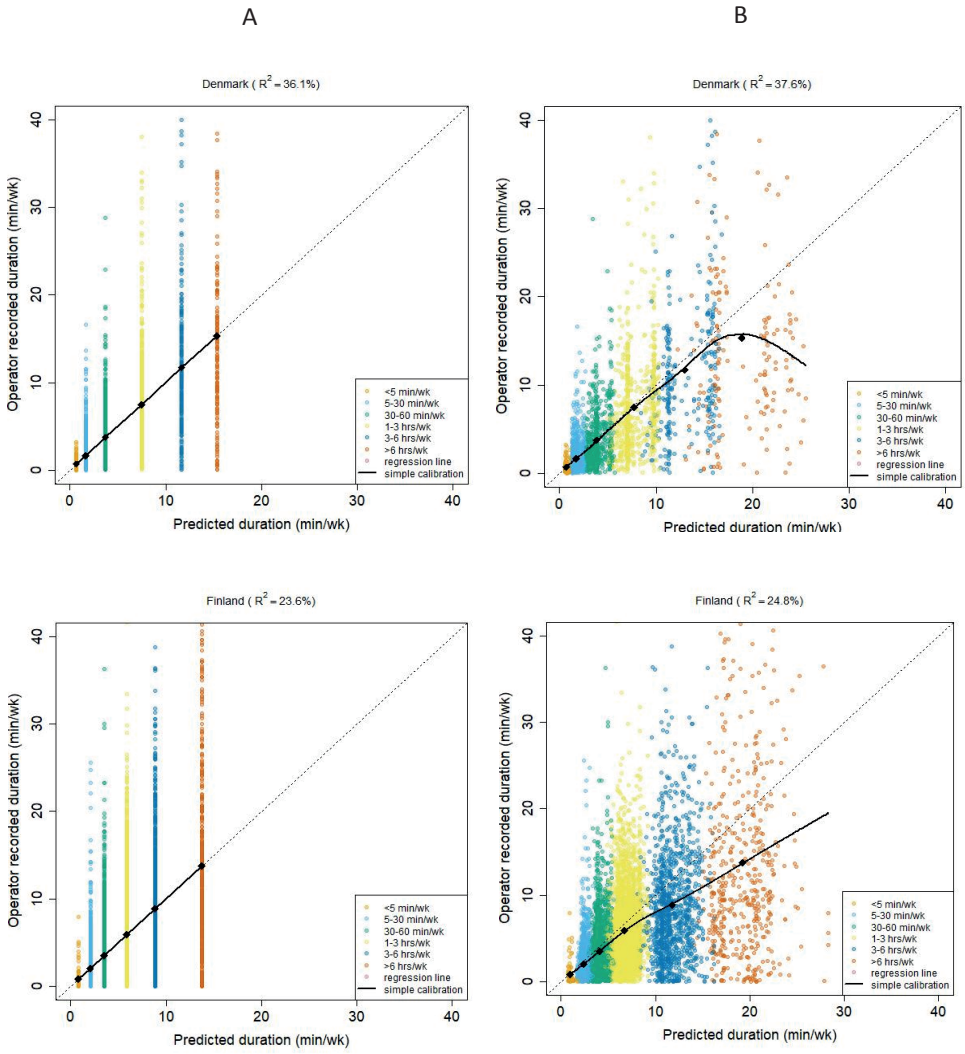
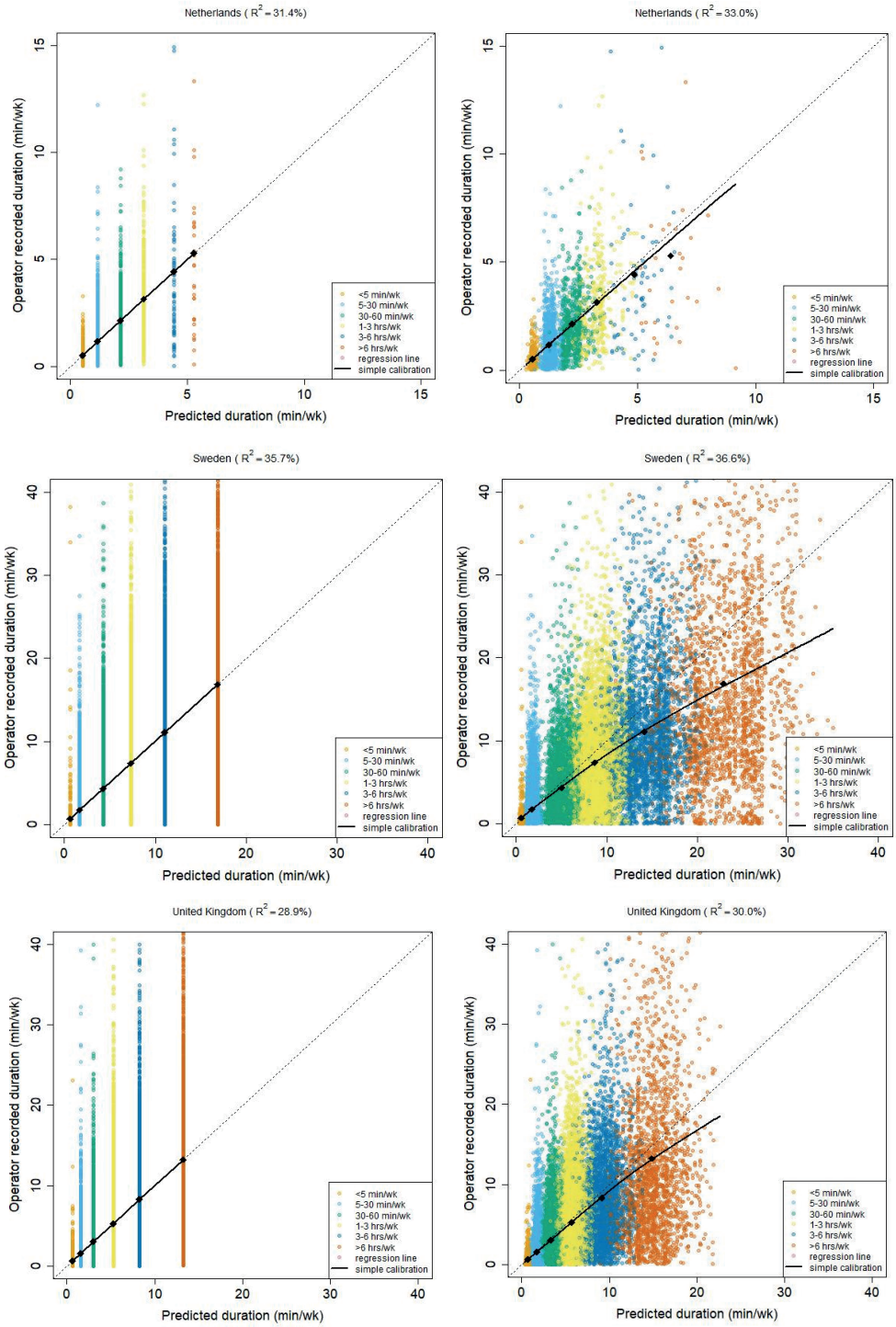

Netherlands
Sweden
United Kingdom
Duration of calls (% of reference)
Age
Underlying variable for self-reported duration of calls
Operator-recorded duration of calls

Figure A3. Observed versus predicted duration of outgoing calls by country for the simple (A) and gamlss-based (B) regression calibration approaches. The regression line was estimated allowing for a non-linear relation between observed and predicted values and allowing the (residual) variance in observed durations to depend on predicted duration using penalized splines (P-splines) as implemented in the gamlss software[8]. Note the different horizontal and vertical scales for the Netherlands.




Netherlands (R2 = 31.4%)
Netherlands (R2 = 33.0%)
Sweden (R2 = 35.7%)
Sweden (R2 = 36.6%)
United Kingdom (R2 = 28.9%)
United Kingdom (R2 = 30.0%)
Operator recorded duration (min/wk)
Predicted duration (min/wk)
<5 min/wk
5-30 min/wk
30-60 min/wk
1-3 hrs/wk
3-6 hrs/wk
>6 hrs/wk
regression line
simple calibration

REFERENCES

1. Carroll RJ, Ruppert D, Stefanski LA, *et al. Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman & Hall/CRC 2006. doi:10.1080/00401706.1997.10485096
2. Freedman LS, Commins JM, Willett W, *et al.* Practice of Epidemiology Evaluation of the 24-Hour Recall as a Reference Instrument for Calibrating Other Self-Report Instruments in Nutritional Cohort Studies : Evidence From the Validation Studies Pooling Project. *Am J Epidemiol* 2018;**186**:73–82. doi:10.1093/aje/kwx039
3. Spiegelman D. Approaches to Uncertainty in Exposure Assessment in Environmental Epidemiology. *Annu Rev Public Health* 2010;**31**:149–63. doi:10.1146/annurev.publhealth.012809.103720
4. Vrijheid M. Validation of short term recall of mobile phone use for the Interphone study. *Occup Environ Med* 2006;**63**:237–43. doi:10.1136/oem.2004.019281
5. Vrijheid M, Armstrong BK, Bédard D, *et al.* Recall bias in the assessment of exposure to mobile phones. *J Expo Sci Environ Epidemiol* 2009;**19**:369–81. doi:10.1038/jes.2008.27
6. Aydin D, Feychting M, Schüz J, *et al.* Impact of random and systematic recall errors and selection bias in case-control studies on mobile phone use and brain tumors in adolescents (CEFALO study). *Bioelectromagnetics* 2011;**32**:396–407. doi:10.1002/bem.20651
7. Samkange-Zeeb F, Berg G, Blettner M. Validation of self-reported cellular phone use. *J Expo Anal Environ Epidemiol* 2004;**14**:245–8. doi:10.1038/sj.jea.7500321
8. Berg G, Schüz J, Samkange-Zeeb F, *et al.* Assessment of radiofrequency exposure from cellular telephone daily use in an epidemiological study: German Validation study of the international case-control study of cancers of the brain--INTERPHONE-Study. *J Expo Anal Environ Epidemiol* 2005;**15**:217–24. doi:10.1038/sj.jea.7500390
9. Heinävaara S, Tokola K, Kurttio P, *et al.* Validation of exposure assessment and assessment of recruitment methods for a prospective cohort study of mobile phone users (COSMOS) in Finland: a pilot study. *Environ Heal* 2011;**10**:14. doi:10.1186/1476-069X-10-14
10. Toledano MB, Auvinen A, Tettamanti G, *et al.* An international prospective cohort study of mobile phone users and health (COSMOS): factors affecting validity of self-reported mobile phone use. *Int J Hyg Environ Health* Published Online First: 2017. doi:https://doi.org/10.1016/j.ijheh.2017.09.008
11. Agogo GO, Van Der Voet H, Van'T Veer P, *et al.* Use of two-part regression calibration model to correct for measurement error in episodically consumed foods in a single-replicate study design: EPIC case study. *PLoS One* 2014;**9**. doi:10.1371/journal.pone.0113160
12. Freedman LS, Schatzkin A, Midthune D, *et al.* Dealing With Dietary Measurement Error in Nutritional Cohort Studies. 2011;**103**. doi:10.1093/jnci/djr189
13. Agogo GO, van der Voet H, van't Veer P, *et al.* Evaluation of a two-part regression calibration to adjust for dietary exposure measurement error in the Cox proportional hazards model: A simulation study. *Biometrical J* 2016;**58**:766–82. doi:10.1002/bimj.201500009
14. Roosbroeck S Van, Li R, Hoek G, *et al.* Traffic-related outdoor air pollution and respiratory symptoms in children: the impact of adjustment for exposure measurement error. *Epidemiology* 2008;**19**:409–16. doi:10.2307/20485657
15. Fraser GE, Stram DO. Regression calibration in studies with correlated variables measured with error. *Am J Epidemiol* 2001;**154**:836–44. doi:10.1093/aje/154.9.836
16. Redmayne M, Smith E, Abramson MJ. A forecasting method to reduce estimation bias in self-reported cell phone data. *J Expo Sci Environ Epidemiol* 2013;**23**:539–44. doi:10.1038/jes.2012.70
17. Petit C, Blangiardo M, Richardson S, *et al.* Association of environmental insecticide exposure and fetal growth with a bayesian model including multiple exposure sources: The PELAGIE mother-child cohort. *Am J Epidemiol* 2012;**175**:1182–90. doi:10.1093/aje/kwr422

18. Bateson TF, Wright JM. Regression calibration for classical exposure measurement error in environmental epidemiology studies using multiple local surrogate exposures. *Am J Epidemiol* 2010;**172**:344–52. doi:10.1093/aje/kwq123
19. Momoli F, Siemiatycki J, McBride ML, *et al.* Probabilistic Multiple-Bias Modeling Applied to the Canadian Data from the Interphone Study of Mobile Phone Use and Risk of Glioma, Meningioma, Acoustic Neuroma, and Parotid Gland Tumors. *Am J Epidemiol* 2017;**186**:885–93. doi:10.1093/aje/kwx157
20. Tokola K, Kurttio P, Salminen T, *et al.* Reducing overestimation in reported mobile phone use associated with epidemiological studies. *Bioelectromagnetics* 2008;**29**:559–63. doi:10.1002/bem.20424
21. Schüz J, Elliott P, Auvinen A, *et al.* An international prospective cohort study of mobile phone users and health (Cosmos): Design considerations and enrolment. *Cancer Epidemiol* 2011;**35**:37–43. doi:10.1016/j.canep.2010.08.001
22. Spiegelman D, Logan R, Grove D. Regression calibration with heteroscedastic error variance. *Int J Biostat* 2011;**7**:4. doi:10.2202/1557-4679.1259
23. Groothuis-oudshoorn K. mice : Multivariate Imputation by Chained. 2011;**45**.
24. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations. *J Stat Softw* 2011;**16**:1–67.https://cran.r-project.org/package=mice
25. Toledano MB, Smith RB, Chang I, *et al.* Cohort profile: UK COSMOS—a UK cohort for study of environment and health. *Int J Epidemiol* 2015;:dyv203. doi:10.1093/ije/dyv203
26. Reedijk M, Lenters V, Slottje P, *et al.* Cohort profile : LIFEWORK , a prospective cohort study on occupational and environmental risk factors and health in the Netherlands. Published Online First: 2018. doi:10.1136/bmjopen-2017-018504
27. Rigby RA, Stasinopoulos DM, Lane PW. Generalized additive models for location, scale and shape. *J R Stat Soc Ser C Appl Stat* 2005;**54**:507–54. doi:10.1111/j.1467-9876.2005.00510.x
28. Plummer M. A program for analysis of Bayesian graphical models using Gibbs sampling. *Proc 3rd Int Work Distrib Stat Comput* 2003;:20–22.https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf (accessed 26 Jun 2018).
29. Spiegelman D, McDermott A, Rosner B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* 1997;**65**:1179S-1186S. doi:10.1093/ajcn/65.4.1179S
30. Blackwell M, Honaker J, King G. A Unified Approach to Measurement Error and Missing Data: Overview and Applications. *Sociol Methods Res* 2017;**46**:303–41. doi:10.1177/0049124115585360
31. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. 1987. doi:10.1002/9780470316696
32. Spiegelman D, Carroll RJ, Kipnis V. Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument. *Stat Med* 2001;**20**:139–60.http://www.ncbi.nlm.nih.gov/pubmed/11135353 (accessed 27 Jun 2018).
33. Wang CY, Hsu L, Feng ZD, *et al.* Regression Calibration in Failure Time Regression. 1997. https://about.jstor.org/terms (accessed 21 Apr 2020).
34. Fraser GE, Stram DO. Regression calibration when foods (measured with error) are the variables of interest: Markedly non-gaussian data with many zeroes. *Am J Epidemiol* 2012;**175**:325–31. doi:10.1093/aje/kwr316
35. Liao X, Spiegelman D, Carroll RJ. Regression calibration is valid when properly applied. Epidemiology. 2013;**24**:466–7. doi:10.1097/EDE.0b013e31828b284b
36. Szpiro AA, Paciorek CJ. Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics* 2013;**24**:501–17. doi:10.1002/env.2233
37. Bartlett JW, Keogh RH. Bayesian correction for covariate measurement error: A frequentist evaluation and comparison with regression calibration. *Stat Methods Med Res* 2018;**27**:1695–708. doi:10.1177/0962280216667764

38. Szpiro AA, Sheppard L, Lumley T. Efficient measurement error correction with spatially misaligned data. *Biostatistics* 2011;**12**:610–23. doi:10.1093/biostatistics/kxq083

39. Vergnaud AC, Aresu M, Kongsgård HW, *et al.* Estimation of TETRA radio use in the Airwave Health Monitoring Study of the British police forces. *Environ Res* 2018;**167**:169–74. doi:10.1016/j.envres.2018.07.015

Chapter 4

Parkinson's disease case ascertainment in prospective cohort studies through combining multiple health information resources

Marije Reedijk, Anke Huss, Robert A. Verheij, Petra H. Peeters, Roel C. H. Vermeulen

Published: PLoS One. 2020;15(7):e0234845. Published 2020 Jul 1.

ABSTRACT

Epidemiological evidence from prospective cohort studies on risk factors of Parkinson's disease (PD) is limited as case ascertainment is challenging due to a lack of registries and the disease course of PD. The objective of this study was to create a case ascertainment method for PD within two prospective Dutch cohorts based on multiple sources of PD information. This method was validated using clinical records from the general practitioners (GPs). Face validity of the case ascertainment was tested for three etiological factors (smoking, sex and family history of PD).

In total 54825 participants were included from the cohorts AMIGO and EPIC-NL. Sources of PD information included self-reported PD, self-reported PD medication, a 9 item screening questionnaire (Tanner), electronical medical records, hospital discharge data and mortality records. Based on these sources we developed a likelihood score with 4 categories (no PD, unlikely PD, possible PD, likely PD). For the different sources of PD information and for the likelihood score we present the agreement with GP-validated cases. Risk of PD for established factors was studied by logistic regression as exact diagnose dates were not always available.

Based on the algorithm, we assigned 346 participants to the likely PD category. GP validation confirmed 67% of these participants in EPIC-NL, but only 12% in AMIGO. PD was confirmed in only 3% of the participants with a possible PD classification. PD case ascertainment by mortality records (91%), EMR ICPC (82%) and self-reported information (62-69%) had the highest confirmation rates. The Tanner PD screening questionnaire had a lower agreement (18%). Risk estimates for smoking, family history and sex using all likely PD cases were comparable to the literature for EPIC-NL, but not for smoking in AMIGO. Using multiple sources of PD evidence in cohorts remains important but challenging as performance of sources varied in validity.

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disease affecting a considerable part of the elderly population, with a prevalence of more than 2% in the population above 65[1,2]. The prevalence of PD will likely increase in the coming years as the population ages and there are currently no disease modifying treatments. PD has some familial forms, however in most cases the disease is idiopathic[3]. Idiopathic PD is thought to be caused by an interaction between aging, genetics, and environmental factors[4]. For example PD is more common among men than women and a lower risk is found among smokers in epidemiological studies[5,6]. Other putative PD risk factors include pesticides, alcohol or coffee consumption and diet[7-10]. Currently many epidemiological studies on PD use a case-control design, due to the gains in efficiency and statistical power, however case-control studies may suffer from reverse causality and other biases such as retrospective recall of lifestyle factors. The lower risk of such biases in prospective cohort studies could make them an important additional resource for studying PD risk factors.

This however, raises an issue as accurate case identification of PD, a relatively rare disease, is important for the validity of effect estimates in epidemiological studies, but currently in many countries no specific registry for PD is available. This makes it a challenge to identify PD cases in prospective studies. Identification is also complicated because PD has a long preclinical period before disease manifests and an unclear moment of onset[7,11]. Further, the diagnostic procedure for PD is symptom-based, its symptoms can also be caused by related conditions such as essential tremor or non-PD parkinsonisms. Screening the total cohort population by a specialist is unfeasible and therefore, PD identification in cohort studies generally rely on medical records and self-reported information[12]. This led Tanner et. al.[13] to develop a series of nine questions on Parkinsonian symptoms in 1990 to help identify PD cases.

The aim of the present paper is to develop a case ascertainment method to identify Parkinson's disease within two prospective cohort studies in the Netherlands (EPIC-NL and AMIGO). PD information was gathered from multiple sources, including the Tanner screening questionnaire[13], self-reported diagnosed PD, self-reported PD medication use, electronic medical records, hospital discharge data and mortality records. The proposed PD algorithm was tested in two manners; first the detected participants were validated against data from the general practitioners (GPs) as in the Netherlands GPs have a complete medical overview of their patients. Second, the algorithm was tested by evaluating three well studied and relatively strong etiological factors: smoking, family history of PD, and sex in this study population, termed in this paper as face validity[14].

MATERIAL AND METHODS

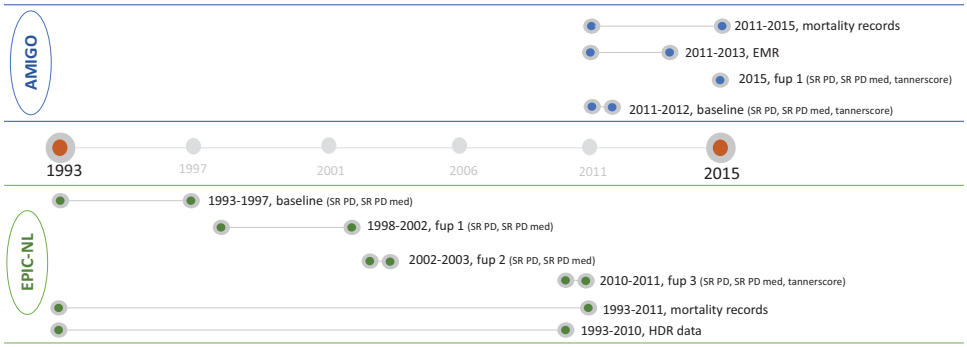
Study population

This study was conducted within the Occupational and Environmental Health Cohort Study (In Dutch: Arbeid, Milieu en Gezondheid Onderzoek, AMIGO) and the European Prospective Investigation into Cancer and Nutrition in the Netherlands (EPIC-NL) cohort[15,16]. The recruitment procedures and study designs of these prospective cohorts are described in more detail elsewhere and in Table S1[15,16]. Briefly, participants in AMIGO were recruited via a national general practitioners (GP) network, ie. the Netherlands Institute for Health Services Research (NIVEL) Primary Care Database in 2011 and 2012. In total 14829 adults, aged 31 to 65 years, from the general population participated. Participants received a baseline questionnaire in 2011/2012 and the first follow-up questionnaire was conducted in 2015 (n=7905, response rate=53%)[16]. Participants from the EPIC-NL cohort were recruited between 1993–1997, either via a breast cancer screening program conducted in the city of Utrecht and neighbouring towns (women aged 49 to 70; EPIC-Prospect), or adults aged 21 to 64 from the general population of three cities, Amsterdam, Maastricht, and Doetinchem (EPIC-Morgen), which together are referred to as EPIC-NL[15]. Participants received a baseline questionnaire between 1993-1997, in which a total of 40011 participant were included[15]. Follow-up questionnaires were conducted in the years 1998-2002 (follow-up 1, n=28022), 2002-2003 in EPIC-Prospect only (follow-up 2, n=12004) and 2010-2011 (follow-up 3, n=13960) (Tables S2 and S3). Participants included in this study signed a written (EPIC-NL) or electronic (AMIGO) informed consent. Approval was obtained from the Institutional Review Board of the University Medical Center Utrecht (EPIC-Prospect, AMIGO) and the Medical Ethical Committee of TNO Nutrition and Food Research (EPIC-Morgen). AMIGO has also been approved according to the governance code of NIVEL Primary Care Database under NZR-00317.005.

Data sources

For the development of the case ascertainment methods two types of data were used: self-reported information from the questionnaires (self-reported PD diagnosis and self-reported PD medication) and registry data (i.e., electronic medical records, hospital discharge registry, mortality records) (Table S4). In Figure 1, the timing of the different sources of PD information is displayed. In total, the follow-up time for EPIC-NL was almost 20 years, for AMIGO it was shorter with a maximum of 5 year for mortality records.

Figure 1. Timeline data collected within AMIGO and EPIC-NL.



SR, self-reported; fup1=follow-up 1; fup2, follow-up 2; fup3, follow-up 3; PD, Parkinson's Disease; med, medication; EMR, electronic medical records; HDR, hospital discharge registry. Green=EPIC-NL, Blue=AMIGO.

Self-reported information

In the questionnaires, self-reported diagnosis of PD and self-reported medication for PD were assessed. The questions were asked closed (yes/no) in the follow-up 3 questionnaire in EPIC-NL and baseline and follow-up questionnaire in AMIGO. We asked for doctor-diagnosed PD in the closed questions. The other questionnaires administered within EPIC-NL had open questions asking for diseases and medication in general.

The open medication question in the first two questionnaires of EPIC-NL were classified in Anatomical Therapeutic Chemical (ATC) codes. We searched for the ATC codes starting with: N04BA, N04BB, N04BC, N04BD and N04BX, which are specific for PD. If ATC codes were not available we identified medication by using substance names drugs and brand names of all PD drugs registered in the Netherlands as search terms. More details can be found in Table S5.

Besides self-reported PD and self-reported PD medication a series of nine questions on Parkinsonian symptoms, such as having a smaller handwriting than before, were assessed in the baseline and follow-up questionnaire in AMIGO and the third follow-up questionnaire in EPIC-NL. This short questionnaire has been developed by Tanner et al. in 1990 and subsequently shown to be predictive for PD diagnosis in case-control studies[13,17,18]. See the full list of questions in Table S6. From these nine questions a score is calculated by summing the number of positive items, resulting in the "Tanner score" (range 0-9). We divided the Tanner score in three categories 0-1 (unlikely), 2-4 (possible), ≥ 5 (probable), based on previous reports[18-21].

Registry data

Hospital discharge registry

The Dutch hospital discharge register (HDR), is coordinated by the Dutch Hospital Association and Order of Medical Specialists[15]. The HDR is a standardized registry of hospital discharge diagnoses. It contains one mandatory principal diagnosis and up to nine additional diagnoses for every hospital discharge in the Netherlands[15]. All diagnoses in the HDR registry are classified according to the International Classification of Disease, ninth version (ICD-9) by medical administrative employees in hospitals[15]. If one of these diagnoses was PD (ICD-9-CM 332), this information was used for case ascertainment in this study. HDR diagnoses were available for EPIC-NL until 31 December 2010. As of 2011 the HDR registry changed and it is not possible to retrieve ICD coded data anymore.

Electronic Medical Records

Electronic Medical Records (EMRs) for AMIGO participants were extracted from the Netherlands Institute for Health Services Research (NIVEL) Primary Care Database[16]. In this database health outcomes are registered by the International Classification of Primary Care-1 (ICPC) and the ATC classification system registered drug prescriptions, which are combined in this study. EMR-based Parkinson's disease is defined with ICPC-code N87. Prescriptions are defined by all ATC codes starting with N04A and N04B. EMR were available for AMIGO from 2011 to 2013.

Mortality registry

Causes of death of deceased were obtained via Statistics Netherlands (CBS). The cause of death register contains up to three causes of death, which are coded according to the ICD-9. If one of the reported diagnoses was PD (ICD-9-CM 332) this information was used for case ascertainment in this study. Data from the cause of death register were available until 31 December 2011 for EPIC-NL and until 31 August 2015 for AMIGO.

Algorithm

Based on the different sources of PD information we developed a probabilistic likelihood score to allocate participants to 4 categories: 0 = no PD, likelihood 1= unlikely PD, likelihood 2 = possible PD, likelihood 3 = likely PD as shown in Table 1. In this score, likelihood category 1 contains Tanner scores two to four. Self-reported diagnosis, self-reported medication and a Tanner score ≥ 5 are in likelihood category 2. Likelihood 3 is a PD diagnosis in the Statistics Netherlands cause of death register, in the HDR or EMR or at least two types of information from the likelihood 2 category. Participants not matching any of these criteria are placed in likelihood category 0.

Table 1. Algorithm for deciding PD likelihood of participants, by available sources of PD information.

Likelihood score 1	Likelihood score 2	Likelihood score 3
Tanner score 2-4 (A1)[a]	Exactly 1 type of evidence from: - Tanner score ≥ 5 (A2)[a] - Self-reported PD diagnosis (B)[b] - Self-reported PD medication (C)[b]	- At least 2 types of evidence from: A2, B and C - At least: PD diagnosis death certificate (D)[c] - A least: PD diagnosis HDR (E)[d] - A least: PD diagnosis EMR (F)[e]

[a] assessed in follow-up 3 questionnaire EPIC-NL, baseline and follow-up AMIGO
[b] assessed in follow up 1 + 2 questionnaires (open questions) and follow-up 3 questionnaires (PD specific) EPIC-NL, and baseline and follow-up AMIGO
[c] assessed from baseline to 31 December 2011 EPIC-NL, baseline to 31 August 2015 AMIGO
[d] assessed from baseline to 31 December 2010 EPIC-NL
[e] assessed from 2011 until 2013 in AMIGO, ATC and ICPC codes
PD, Parkinson's Disease; EMR, electronic medical records; HDR, hospital discharge registry; ATC, Anatomical Therapeutic Chemical; ICPC, International Classification of Primary Care.

GP verification

The algorithm was verified using questionnaires collected between 2015 and 2017 at the GPs of participants. In the Netherlands, GPs have a complete overview of the medical status of their clients listed in their practices as they receive information back from any other health services such as specialists in hospitals. Virtually all Dutch citizens are listed with a GP as it is obligatory in the Netherlands to have a GP. The GP of the participants was contacted if participants gave informed consent and information on current GP was available. All GPs of participants in likelihood 3 were contacted. Due to restrictions on time and resources, a subsample of PD likelihood 1 and 2 were included in the validation study.

EPIC-NL

In EPIC-NL the GP validation was performed for all participants with likelihood 3 (n=176) and all participants with self-reported PD diagnosis or medication data in likelihood 2 (n=46). Additionally, a random 100 participants each from likelihood 1 and 2 were selected for GP validation (see Table S7).

AMIGO

In AMIGO all participants with likelihood 3 (n=168) were included in the GP validation study. From likelihood 2, all participants with self-reported data were selected for GP validation (n=3) and additionally 100 randomly selected participants from likelihood 2 based on the Tanner score (Table S7). PD is listed as a chronic disease in the EMR system and ones in the system it will remain in the system although diagnosis might be changed.

Procedure

In both EPIC-NL and AMIGO, general practitioners of the participants included in the validation study were contacted by mail and given a short questionnaire with eight questions. Also the GP of deceased participants were contacted as most often information in the system is still available. We largely followed the same procedure as the NeuroEPIC4PD study which has been described in detail by Gallo et al.[22]. An allowance equal to the price of a GP consult of less than 20 minutes was given to the GP per questionnaire that was returned (~9 euro). Non-responding GPs received a reminder by mail and a new copy of the questionnaire. After this reminder the non-responding GPs were contacted by telephone to answer a limited set of questions (e.g. Parkinson diagnosis and year of diagnosis).

QUESTIONNAIRE GP VALIDATION

The most important data retrieved from the GP questionnaire was PD diagnosis and the year of diagnosis. The other questions were on the name and address of a possible new GP of the participant, name and address of the treating neurologist, PD surgery, PD medication use, PD symptoms and differential diagnosis (other forms of Parkinsonism and essential tremor).

Face validity

We further evaluated the performance of the algorithm by performing logistic regression with known etiological factors e.g. baseline self-reported smoking (divided into never, ever, current), sex and self-reported family history of PD.

Statistical analyses

From the GP validation exercise, we calculated the number and percentage of GP-confirmed cases, participants reported not to have PD and participants with a differential diagnosis (tremor by other cause or other Parkinsonism). The number and percentage GP-confirmed cases was also calculated for each separate source of PD information (i.e. self-reported PD diagnosis, self-reported PD medication, Tanner score, mortality records, HDR and EMR).

Logistic regression models were used to calculate odds ratios for known PD risk factors (first degree family history of PD, baseline smoking, sex and age). For likelihood 3, all participants with likelihood 2, 1 and 0 were taken as controls. Analyses were corrected for baseline age (continuous), education (low, medium, high), sex and cohort (in combined analyses). We also performed analyses without the variable education. We conducted the following sensitivity analyses to assess effects of possible outcome misclassification on risk estimates: 1) for PD likelihood 3, participants with likelihood 1 and 0 were taken as controls, as there may be some hidden PD cases in the second likelihood, 2) for PD

likelihood 3, participants with likelihood 0 were taken as controls, as there may be some hidden PD cases in the second and first likelihood score, 3) GP validated PD cases only, with participants with likelihood 0,1 and 2 taken as controls. We also stratified analyses for cohort, age and sex. All statistical analyses were conducted using SAS, version 9.4 (SAS Institute Inc., Cary, North Carolina, USA) and R statistical software (version 3.4.3).

RESULTS

Study population descriptives and PD evidences

Baseline characteristics of the combined cohort, and for the sub cohorts AMIGO and EPIC-NL separately, are shown in Tables 2 and S8. The main differences between the two cohorts are due to differences in recruitment procedures. The higher proportion of female participants within EPIC-NL is

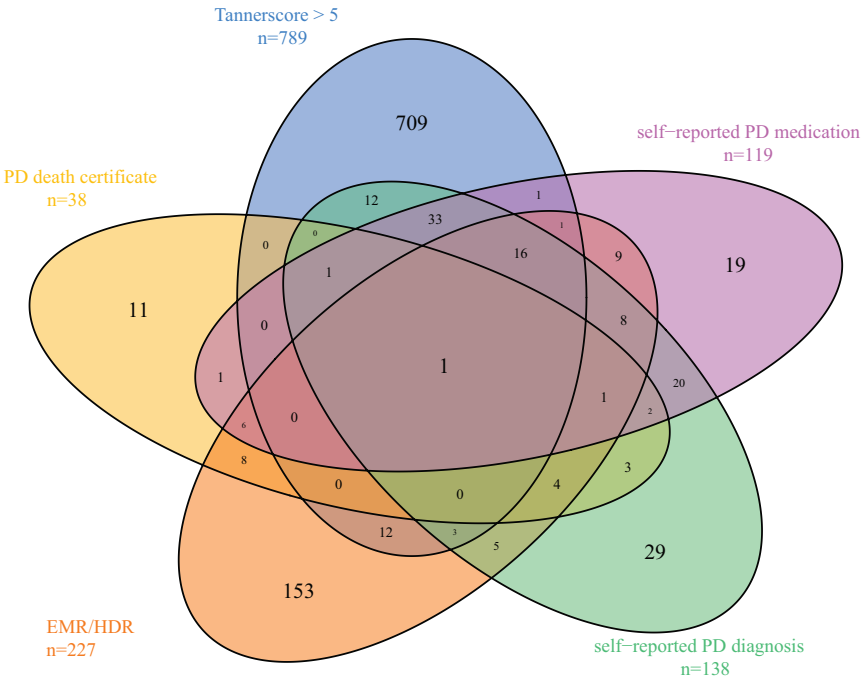
Table 2. Baseline characteristics and percentage of participants with a source of PD information according to cohort (AMIGO, EPIC-NL, Combined).

	AMIGO	**EPIC-NL**	**Combined**
Total N	14829	40011	54825
Age (median, IQR)	51.0 (43.0-59.0)	51.4 (41.9-57.6)	51.4 (42.3-58.0)
Female (%)	8268 (55.8%)	29751 (74.4%)	38007 (69.3%)
Past smoker (%)	5744 (38.8%)	12440 (31.2%)	18179 (33.3%)
Current smoker (%)	2322 (15.7%)	12164 (30.5%)	14484 (26.5%)
Questionnaire data:			
Self-reported PD medication use (%)	32 (0.2%)	87 (0.2%)	119 (0.2%)
Self-reported PD doctor diagnosis (%)	26 (0.2%)	112 (0.3%)	138 (0.3%)
Tanner score 2-4 (%)	1874 (12.6%)	3068 (7.7%)	4939 (9.0%)
Tanner score ≥ 5 (%)	242 (1.6%)	547 (1.4%)	789 (1.4%)
Tanner score 2011 (mean, sd)	0.60 (1.12)	1.03 (1.48)	0.81(1.33)
Tanner score 2015 (mean, sd)	0.54 (1.05)	-	0.54 (1.05)
Registry data:			
HDR PD diagnosis (%)	-	94 (0.2%)	94 (0.2%)
EMR PD (%)	133 (0.9%)	-	133 (0.2%)
EMR PD ICPC diagnosis (%)	21 (0.1%)	-	21 (0.1%)
EMR PD ATC medication (%)	130 (0.9%)	-	130 (0.2%)
PD on death certificate (%)	1 (0.0%)	37 (0.1%)	38 (0.1%)

Participants may have multiple sources of PD information.
Age, Age at study entry; PD, Parkinson's Disease; EMR, electronic medical records; HDR, hospital discharge registry; SD, standard deviation; IQR, Interquartile range; N, sample size; ATC, Anatomical Therapeutic Chemical; ICPC, International Classification of Primary Care.

due to the recruitment via the breast cancer screening program in EPIC-Prospect. AMIGO participants are more often higher educated and less often current smokers compared to EPIC-NL participants. The number of participants for each information source of PD is shown in Table 2, which shows that after the Tanner score, HDR and EMR are the most frequent PD information sources. Figure 2 shows the overlap between the different PD information sources for the combined cohort which indicates that participants had often only one information source of PD for the Tanner score ≥ 5 (89%) or EMR/HDR (72%). Supplementary figures 1 and 2 display the overlap for AMIGO and EPIC-NL separately which looked similar. In Table 3 the number of participants assigned to each of the four PD category is shown. In total 346 out of 54825 (0.63%) participants were assigned the highest PD likelihood.

Figure 2. Venndiagram of different sources of PD information in the Combined cohort.



PD, Parkinson's Disease; EMR, electronic medical records; HDR, hospital discharge registry.

Table 3. Frequency of PD likelihood scores in AMIGO, EPIC-NL and Combined cohort.

PD Likelihood	Cohort	Number (%)
0	AMIGO	12193 (82.2%)
	EPIC-NL	36260 (90.6%)
	Combined	48442 (88.4%)
1	AMIGO	2223 (15.0%)
	EPIC-NL	3039 (7.6%)
	Combined	5258 (9.6%)
2	AMIGO	243 (1.6%)
	EPIC-NL	536 (1.4%)
	Combined	779 (1.4%)
3	AMIGO	170 (1.2%)
	EPIC-NL	176 (0.4%)
	Combined	346 (0.6%)

GP validation of the PD likelihood score

A sample of 668 participants out of 6383 participants with a PD likelihood score of 1 or higher were selected for GP validation (10%). For 167 participants, there was no information on PD available from the GP either by a lack of information the GP had available (n=125) or because there was no response from the GP (n=42). Table 4 provides an overview of the number of participants that were confirmed having PD by their GP. Of the 501 participants with information on PD diagnosis, 85 were confirmed to have PD (17%). None of the likelihood 1 participants had a PD diagnosis confirmed by their GPs. In total 5 (3%) cases of likelihood 2 were confirmed having PD and 14 (7%) participants were reported to have tremor by another cause than PD or another form of Parkinsonism. 34% of the participants with a GP questionnaire returned in likelihood 3 were confirmed to have a PD diagnosis. There were large differences between AMIGO and EPIC-NL, as respectively 12% and 67% of likelihood 3 participants were confirmed by their GP. PD confirmed cases by their GP were older and had more often family with a history of PD (Table S9) compared to non-confirmed cases.

Table 4. Validation of PD Likelihood scores by General Practitioner for EPIC-NL, AMIGO and Combined cohort.

EPIC-NL						
PD Likelihood	Participants selected for validation	Returned GP questionnaire (% of participants selected)	Information available diagnosis (% of participants selected)	PD diagnosis confirmed by GP (%)	No PD diagnosis by GP (%)	Other Parkinsonism/ tremor by other cause (%)
1	99	96 (97%)	72 (73%)	0 (0%)	72 (100%)	2 (3%)
2	141	135 (96%)	110 (78%)	5 (5%)	105 (95%)	6 (5%)
3	160	152 (95%)	93 (58%)	62 (67%)	31(33%)	8 (9%)
AMIGO						
PD Likelihood	Participants selected for validation	Returned GP questionnaire (% of participants selected)	Information available on PD diagnosis (% of participants selected)	PD diagnosis confirmed by GP (%)	No PD diagnosis by GP(%)	Other Parkinsonism/ tremor by other cause (%)
2	100	92 (92%)	81 (81%)	0 (0%)	81 (100%)	8 (10%)
3	168	153 (91%)	145 (86%)	18 (12%)	127 (88%)	15 (10%)
Combined						
PD Likelihood	Participants selected for validation	Returned GP questionnaire (% of participants selected)	Information available on PD diagnosis (% of participants selected)	PD diagnosis confirmed by GP (%)	No PD diagnosis by GP (%)	Other Parkinsonism/ tremor by other cause (%)
1	99	96 (97%)	72 (73%)	0 (0%)	72 (100%)	2 (3%)
2	241	227 (94%)	191 (79%)	5 (3%)	186 (97%)	14 (7%)
3	328	305 (93%)	238 (73%)	80 (34%)	158 (66%)	23 (10%)

PD, Parkinson's Disease; GP, general practitioner

Validation of PD information sources

A PD diagnosis on a death certificate was found in 10 cases (90.9%) to correspond with a GP-confirmed diagnosis, but only 11.7% of all PD cases were identified on the basis of a death certificate (Tables 5 and S10). The agreement was also high for self-reported PD diagnosis (62.4%) and self-reported PD medication (68.5%). The agreement for EMR was 14% with large differences between EMR based on ATC medication codes (12.6%) or ICPC diagnosis code (82.4%). The mean Tanner score was higher for GP validated cases than non-confirmed cases (Table S10). The mean age at diagnosis for validated GP cases was 59 years (range 34 to 68) in AMIGO compared to 68 years (range 48 to 85 year) in EPIC-NL (Table S9).

Table 5. Verification of different sources of PD information by information retrieved from the General Practitioners for AMIGO, EPIC-NL and Combined cohort.

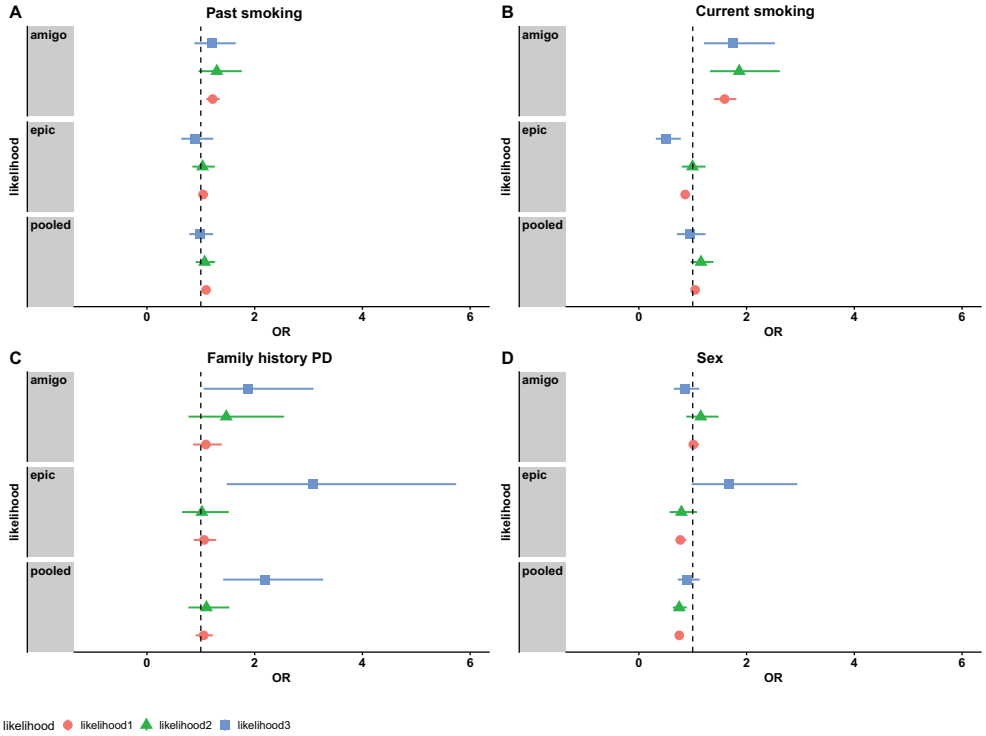
Information source PD	N evidence available	N positive evidence	%	95%CI	Confirmed PD status at GP		
					PD	%	95%CI
AMIGO	226				18		
PD diagnosis (SR)	226	21	9.3	5.8-13.9	15	71.4	47.8-88.7
PD medication(SR)	226	26	11.5	7.7-16.4	15	57.7	36.9-76.6
Tannerscore 2011 ≥ 2	226	162	71.7	65.3-77.5	15	9.3	5.3-14.8
Tannerscore 2011 ≥ 5	226	108	47.8	41.1-54.5	10	9.3	4.5-16.4
PD EMR	226	114	50.4	43.7-57.1	16	14.0	8.2-21.8
PD EMR ICPC diagnosis	226	17	7.5	4.4-11.8	14	82.4	56.6-96.2
PD EMR ATC medication	226	111	49.1	42.4-55.8	14	12.6	7.1-20.3
PD on death certificate	226	1	0.4	0.0-2.4	1	100	2.5-100
EPIC-NL	275				67		
PD diagnosis (SR)	275	72	26.2	21.1-31.8	43	59.7	47.5-71.1
PD medication(SR)	275	47	17.1	12.8-22.1	35	74.5	59.7-86.1
Tannerscore 2011 ≥ 2	233	211	90.6	86.1-94.0	36	17.1	12.2-22.8
Tannerscore 2011 ≥ 5	233	116	49.8	43.2-56.4	31	26.7	18.9-35.7
PD HDR	275	42	15.3	11.2-20.1	23	54.8	38.7-70.2
PD on death certificate	275	10	3.6	1.8-6.6	9	90.0	55.5-99.7
COMBINED	501				85		
PD diagnosis (SR)	501	93	18.6	15.3-22.2	58	62.4	51.7-72.2
PD medication(SR)	501	73	14.6	11.6-18.0	50	68.5	56.6-78.9
Tannerscore 2011 ≥ 2	459	373	81.3	77.4-84.7	52	13.9	10.6-17.9
Tannerscore 2011 ≥ 5	459	224	48.8	44.1-53.5	41	18.3	13.5-24.0
PD on death certificate	501	11	2.2	1.1-3.9	10	90.9	58.7-99.8

SR, self-reported; PD, Parkinson's Disease; EMR, electronic medical records; HDR, hospital discharge registry; GP, general practitioner; ATC, Anatomical Therapeutic Chemical; ICPC, International Classification of Primary Care.

Association with known etiological factors - face validity

Characteristics for likelihood 3 compared to likelihood 0-2 are shown in Tables S11 and S12. The results of the face validity for smoking, sex and family history for PD likelihood scores 1-3 compared to likelihood 0 are shown in Figure 3 and Tables S13 and S14 for likelihood 3. Current smoking (at baseline) as compared to never smoked showed an OR of 0.88(95% CI: 0.65-1.18) for likelihood 3 (compared to likelihood 0-2). Similarly, for past smokers, only for likelihood 3 risk estimates below unity were observed, compared to never smokers. Large differences existed between the two cohorts for past and current smoking, the estimates found in EPIC-NL were noticeably lower. First degree family history of PD increased the risk of PD for likelihood 3 with an OR of 2.29 (95% CI: 1.43-3.49) in the Combined cohort (Table S13). PD risk of sex (male vs. female) provided an OR of 0.79 (95% CI: 0.57-1.07) for AMIGO. The cohort specific odds ratio in EPIC-NL was 1.56 (95%CI: 0.90-2.76) for male vs female.

Figure 3. Adjusted logistic regression by PD likelihood scores for known PD risk factors past smoking (baseline)(A), current smoking (baseline)(B), 1st degree family history of PD (C), and sex (D) in AMIGO, EPIC-NL and Combined cohort.



Reference group for smoking were never smokers. PD, Parkinson Disease; OR, odds ratio. Likelihood 3 compared to likelihood 0. Likelihood 2 compared to likelihood 0. Likelihood 1 compared to likelihood 0.
* Adjusted for age at baseline, baseline educational level, sex and cohort (combined analyses)

Sensitivity analyses with likelihood 0-1 and likelihood 0 instead of likelihood 0-2 as controls for likelihood 3 showed very similar results (Table S14). The results of the face validity analyses for smoking, sex and PD family history for the GP confirmed PD cases (n=85) was similar to the results of likelihood 3 cases (n=346) except for sex, although statistical power was low (Table S15). Sensitivity analyses without EMR ATC medication in likelihood 3 for AMIGO showed that the sex association reversed to 1.65 (95%CI: 1.02-2.71) compared to 0.81(95%CI: 0.59-1.10) with EMR medication in the highest likelihood (Table S16). Stratified analysis by sex showed higher ORs for PD family history for females (Table S17). Stratification by age showed differences for smoking and PD family history but these findings were not consistent in the two cohorts (Table S17).

DISCUSSION

In this paper, we combined multiple sources of health information regarding Parkinson's disease to assign participants a probabilistic likelihood score for PD. We validated this algorithm against information retrieved from general practitioners and computed risk estimates for three well-established risk factors (smoking, sex and family history) of PD to judge whether they were in line with expectation.

GP validation

In our study 0.63% of the participants were assigned to the highest likelihood score which is lower than the 1.4% expected prevalence in the general population above 55[23,24]. If we combine self-reported PD and self-reported PD medication with likelihood 3 (n=394, 0.72%) the prevalence in our cohort slightly increases. A possible explanation for the lower prevalence is that in EPIC-NL a healthy volunteer effect has been observed[25] and our study population (especially in the AMIGO cohort) is younger. Besides there is a possibility of unidentified PD cases in the lower likelihood scores.

In total 85 participants were diagnosed with PD by their general practitioner. We were unable to verify the PD status of 167 participants invited for GP validation. From likelihood 2 a modest 3% of the participants were confirmed to have PD. The highest likelihood category had a verification rate of 34%. Verification rates for AMIGO and EPIC-NL differed considerably as the agreement for EPIC-NL participants was 67% and that of AMIGO was 12% for likelihood 3. In addition to comparing the PD likelihood score to GP clinical records we also compared the different sources of PD information against GP information. Self-reported PD medication had a high agreement as 69% were confirmed by their GP. Mortality records from the Statistics Netherlands cause of death registry only found 10 of the 85 GP confirmed cases (12%) but the ones identified had a high likelihood of having PD (91%). Two other registries used in this study were the EMR and the HDR in AMIGO and EPIC-NL respectively. A positive PD diagnosis in the EMR had a much lower agreement with GP validated cases than the HDR (14% compared to 55%). The EMR system found 16 out of the 18 PD validated cases in AMIGO, but also 98 participants were identified as having PD by the EMR system but not by the GP. This could be caused by the fact that if a GP suspects PD in a patient, it is noted in the EMR system where it remains as PD is regarded as a chronic disease. Alternatively, PD medication is sometimes used as a diagnostic tool for PD which may have led to a higher amount of false positives in AMIGO. This is supported by the lower agreement found for ATC medication codes in comparison to ICPC diagnosis code which showed a high agreement with the GP records (82%).

PD likelihood scores 1 and 2 were largely assigned because of the reporting of PD symptoms via the screening questionnaire of Tanner. There were only 5 participants with likelihood score 2 that

were confirmed to have PD by their GP, so Tanner by itself was not predictive for PD in a prospective setting. It is possible that after a longer period of follow-up the verification rates will increase since the screenings questionnaire was administered only a couple of years (±4) before the end of follow-up in both EPIC-NL and AMIGO.

Face validity of likelihood algorithm

The effect estimates of three well-known risk factors (smoking, sex, and family history of PD) were similar to what we expected based on the literature for the highest likelihood in EPIC-NL. However, estimates were attenuated towards the null for smoking and reversed for sex in AMIGO. A possible explanation for the attenuated finding for smoking in AMIGO is that ever or currently smoking participants may have had more medical visits due to smoking-related medical conditions, and this provides more opportunity for PD to be identified. The PD risk of first degree family history was OR 2.29 (95% CI: 1.43-3.49) for likelihood 3 in the Combined cohort and was slightly smaller than the odds ratio found in the meta-analysis of Noyce et. al.[14]. Males had comparable PD risk to women in our study, while the previous literature showed higher risk of PD for males. However, the estimate of EPIC-NL (OR: 1.56, 95% CI: 0.90-2.76) was comparable with the OR found in a systemic review on sex and PD from 2004[26]. A reversed effect estimate was found for sex in AMIGO. However, if we eliminated cases identified based on EMR medication the effect estimate for sex was in the expected direction and of similar magnitude (Table S17). In our study, different effects estimates were found for the two different cohorts. The effect estimates found in EPIC-NL for the face validity, which were in the expected direction, and the high agreement of likelihood 3 by the GP in EPIC-NL indicate that the algorithm worked well in identifying PD cases. However the effect estimates of AMIGO were attenuated or even reversed for sex of what we expected and likelihood 3 had a much lower agreement with GP validated cases (12%). This means that the proposed algorithm was less capable in identifying PD cases in AMIGO. A possible explanation for this is the performance of the EMR based on ATC medication codes. 133 participants had a positive result in the EMR registry based on ATC codes or ICPC diagnosis, which was an important source to end up in the highest likelihood in AMIGO. The EMR registry on ATC medication codes had an agreement of only 13%. Therefore, ATC medication codes should be reconsidered in future use of the PD algorithm for AMIGO. The risk of PD increases with age, which was also seen in our study. Baseline age was comparable in both cohorts but as EPIC-NL had 15 years longer follow-up time, the age of EPIC-NL participants at the end of follow-up was higher. This is not reflected in the tables (results section) because only age at baseline is reported.

The past couple of years more effort is conducted on ascertainment of PD in population-based cohort studies[12,22,27,28]. These studies were also conducted originally for other health outcomes, such as the Cardiovascular Health Study[12,28], and Framingham Heart Study[27]. They indicate that using multiple

sources of PD information is important but also that self-reported data remains important in cohort studies as medical registry data have limitations and on average had lower agreements, similar as our study[27,28].

Strengths and limitations

For PD ascertainment we used multiple sources of PD information retrieved at multiple time points during on-going prospective studies which was a strength as disease misclassification can be reduced by multiple sources[28-30]. Using only one type of evidence might cause a proportion of the cases not to be identified[2,5]. Another strength was the long follow-up period of more than 20 year in EPIC-NL. A long follow-up period is important because of the long preclinical period of PD[31]. Another strength was the large sample size, making this one of the largest cohort studies on Parkinson's disease in the Netherlands. Our study was limited by the fact that we only validated a part of our study population and not all GPs responded to our questionnaire. We only selected 668 participants out of the total study population of 54825 for GP validation and therefore we were not able to calculate population prevalence, sensitivity, specificity, positive and negative predictive values or c-statistics. Second, by validating only this subsample there might be unidentified cases of PD that were not included in the sub-sample. However, for a relatively rare disease like PD, unidentified cases do not have much influence on the risk estimates in a cohort study as the effect of these participants is diluted because of a large surplus of true controls[32]. Nevertheless, the PD cases that are identified must be representative of all cases as to not bias risk estimates. Another weakness of the case ascertainment method regarding disease misclassification is the possibility of ascertaining cases that are not truly PD cases. Unfortunately, verification of PD status by a neurologist, the golden standard in PD research, was not possible in this study setup. We verified the algorithm against information retrieved from the GPs, which is after neurological examination by a physician, the best way of validation in the Netherlands because GPs are regarded to have a complete overview of the medical status of patients including information from the treating neurologist. However, there could be delays and omissions in the GP system.

Conclusion

We applied a case ascertainment strategy for PD in two prospective cohort studies, which generated a probabilistic likelihood score to classify participants into four categories. The highest likelihood performed reasonably well when comparing the obtained effect estimates of known PD risk factors with the literature, particular in EPIC-NL. The case ascertainment algorithm worked well to identify participants with PD in EPIC-NL as likelihood 3 had a GP confirmation rate of 67%. The AMIGO algorithm can be improved by not incorporating recorded PD medication use which generated most

false positives. Overall, self-reported information and PD evidence in the mortality registry performed well with high agreements (62%-91%). Other sources of PD information gave varying results and performance is dependent on the source and underlying cohort demographics.

Acknowledgments

We acknowledge the contribution of dr. J.M.A. Boer and the National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands, for their contribution and ongoing support to the EPIC-NL Study. The authors thank Susan van Hemert from the University Medical Center Utrecht, Jan van der Laan from the Dutch National Institute for Public Health and the Environment (RIVM) and Inka Pieterson from Utrecht University for helping with data management tasks. They also are grateful to Joke Metselaar (Julius Center), Ingrid Dahmen (Utrecht University), Guido Roos (Utrecht University) and Leandro Cortellini (Utrecht University) for their contribution to the data collection.

SUPPLEMENTARY TABLES AND FIGURES CHAPTER 4

Table S1. Cohort characteristics of EPIC-NL (EPIC-MORGEN,EPIC-PROSPECT) and AMIGO.

Cohort characteristics	EPIC-MORGEN	EPIC-PROSPECT	AMIGO
Type of participants and source population	General population sample of men and women age 20-59 at recruitment	Women aged 49–70 at recruitment, who received a call-up for the Dutch breast cancer screening	General population sample of men and women age 31-65 at recruitment
Place of recruitment	Cities of Amsterdam, Doetinchem and Maastricht	City of Utrecht and its vicinity	Whole of the Netherlands
Participation rate at baseline	44.6%	34.5%	16%
Number of participants	22,654	17,357	14,829
Time of recruitment	1993-1997	1993-1997	2011-2012
Procedure of recruitment	Sample of men and women from the general population and included those who agreed when asked for participation.	Women selected (by age) for breast cancer screening were asked for participation by filling in two questionnaires. Those that returned the questionnaires were asked to participate.	Men and women recruited a general practitioners network, with a maximum of one person per household.

Table S2. Timing of questionnaires of EPIC-PROSPECT, EPIC-MORGEN, AMIGO.

Questionnaire	EPIC-MORGEN	EPIC-PROSPECT	AMIGO
Baseline	1993-1997	1993-1997	2011-2012
Follow-up 1	1998-2002	1998-2000	2015
Follow-up 2	2003-2007 (Doetinchem only)	2002-2003	
Follow-up 3	2010-2011 (Amsterdam + Maastricht)	2011	

Table S3. Questionnaire response rates in EPIC-NL (EPIC-PROSPECT, EPIC-MORGEN) and AMIGO

	AMIGO		EPIC-PROSPECT		EPIC-MORGEN		EPIC-NL	
	No. of participants	*% of baseline*	*No. of participants*	*% of baseline*	*No. of participants*	*% of baseline*	*No. of participants*	*% of baseline*
Baseline	14829		17357	-	22655	-	40011	-
FU1	7905	53%	13201	76.1 %	14821	65.4%	28022	70.0%
FU2			12004	69.0%	-	-	-	-
FU3			7784	44.8%	6176	27.3%	13960	34.9%

Table S4. Different sources of PD information in EPIC-NL and AMIGO

Evidence	EPIC-NL	AMIGO
Self-reported PD diagnosis	Yes: baseline, follow-ups	Yes: baseline, follow-up
Self-reported PD medication	Yes: baseline, follow-ups	Yes: Baseline, follow-up
Tanner questionnaire	Yes: follow-up 3	Yes: Baseline, follow-up
HDR	Yes: until 31 December 2010	No
EMR	No	Yes: from 2011 till 2013
Mortality	Yes: until 31 December 2011	Yes: until 31 December 2015

PD, Parkinson's Disease; EMR, electronic medical records; HDR, hospital discharge registry.

Table S5. Substance names and brand names used as search terms in EPIC-NL

ADARTREL
AKINETON
ALMIRIDE
AMANTADINE
APOBENZTROP
APOGO
APOKYN
APOMORFINE HYDROCHLORIDE
APOTRIHEX
APRICOLIN
ARNALEVOCAP
ARTANE
ATURBAN
AZILECT
BENSERAZIDE
BENZATROPINE
BENZHEXOL
BENZTROP
BIPERIDEEN
BORNAPRINE
BROMOCRIPTINE
BUDIPINE
CABASER
CABERGOLINE
CABERLIN
CARBIDOPA
CLARIUM
CO-BENELDOPA
CO-CARELDOPA
COGENTIN
COMTAN
COMTESS
CORBILTA
CRIPAR
CYCLOSET
DACEPTON
DAQUIRAN
DEKINET
DEPRENALINE
DEPRENIL
DEPRENYL
DEXETIMIDE
DIHYDROERGOCRYPTINE
DISIPAL
DISIPALETTE
DOPERGIN
DOSTINEX
DOSTINEX
DUODOPA
ELDEPRIL
ELDEPRYL
ENCAPIA
ENTACAPON
ENTACAPONE
ETANAUTINE
GLEPARK
IXENSE
JUMEX
KEMADRIN
L-DOPA
LEGANTO
LEPTICUR
LEVODOPA
LEVODOPUM
LISURIDE
MADOPAR
MELEVODOPA
METIXENE
MIRAPEX
MIRAPEXIN
MODOPAR
NAUTAMINE
NEUPRO
OPRYMEA
ORFENADRINE
ORPHENADRINE
PACITANE
PARA LEST
PARKIN
PARLODEL
PARSIDAN
PARSIDOL
PERGOLIDE
PERGOLIDEMESILAAT
PERMAX
PHENGLUTARIMIDE
PIRIBEDIL
PLAMIPEXOL
PRAMIPEXOLE
PRAMITHON
PROCLACAM
PROCYCLIDINE
PROFENAMINE
PROLOPA
PRONORAN
RASAGILINE
REPREVE
REQUIP
REVANIL
RONIROL
ROPINIROL
ROTIGOTINE
SELEGILINE
SINEMET
SIFROL
SORMODREN
SPONTANE
STALEVO
SYMMETREL
TASMAR
TOLCAPON
TRASTAL
TREMARIL
TREMBLEX
TRIHEX
TRIHEXYFENIDYL
TRIVASTAL
TRIVASTAN
TROPATEPINE
UPRIMA
ZELAPAR

Table S6. Questions of the Tanner Questionnaire in AMIGO and EPIC-NL

No.	Question [a]
1.	Do you have trouble arising from a chair?
2.	Is your handwriting smaller than it once was?
3.	Do people tell you that your voice is softer than it once was?
4.	Is your balance poor?
5.	Do your feet ever seem to get stuck to the floor?
6.	Do people tell you that your face seems less expressive than it once did?
7.	Do your arms and legs shake?
8.	Do you have trouble buttoning buttons?
9.	Do you shuffle your feet and/or take tiny steps when you walk

[a] A Dutch translation of the questionnaire was used.

Table S7. Frequency of participants per likelihood score selected for GP follow-up in EPIC-NL and AMIGO

	EPIC-NL		**AMIGO**	
Likelihood	Frequency	Selected for GP follow-up	Frequency	Selected for GP follow-up
0	36261	0	12193	0
1	3040	100 (random)	2223	0
2: TS	490	100 (random)	240	100 (random)
2: other*	46	46	3	3
3	176	176	170	218

* self-reported Parkinson disease or self-reported medication.
TS, Tanner questionnaire Score; GP, general practitioner

Table S8. Baseline characteristics of AMIGO, EPIC-NL and Combined Cohort.

	EPIC-NL	AMIGO	COMBINED
Number of participants (%)	40011	14829	54825
Age at baseline			
Mean (SD)	49.21(11.90)	50.65(9.37)	49.60(11.29)
Sex (%)			
Male	10260(25.6%)	6561(44.2%)	16818(30.7%)
Female	29751(74.4%)	8268(55.8%)	38007(69.3%)
Education (%)			
Low	24198(61.0%)	4537(30.6%)	28731 (52.7%)
Medium	7407(18.7%)	4627(31.2%)	12033(22.1%)
High	8095(20.4%)	5656(38.2%)	13741(25.2%)
Missing	311	9	320
Smoking status at baseline (%)			
Never smoker	15243 (38.3%)	6740(45.5%)	21975(40.2%)
Past smoker	12440(31.2%)	5744(38.8%)	18179(33.3%)
Current smoker	12164(30.5%)	2322(15.7%)	14484(26.5%)
Missing	164	23	187
Family history PD 1st degree (%)*			
Yes	638(4.6%)	496(3.3%)	1133(3.9%)

*Only available for follow-up 3 in EPIC-NL, % calculated based on these participants.
PD, Parkinson's Disease; SD, standard deviation

Table S9. Baseline characteristics PD versus no PD validated by GP for combined cohort.

	PD cases	No PD cases	p-value
Number of participants (%)	85	416	
Age at baseline			
Mean (SD)	59.52(6.99)	53.81(9.13)	<0.001
Sex (%)			
Male	22(25.9%)	116(27.9%)	0.808
Female	63(74.1%)	300(72.1%)	
Education (%)			
Low	53(63.1%)	240(58.5%)	0.233
Medium	10(11.9%)	81(19.8%)	
High	21(25.0%)	89(21.7%)	
Missing	1	6	
Smoking status at baseline (%)			
Never smoker	49(57.6%)	165(40.0%)	0.003
Past smoker	29(34.1%)	159(38.6%)	
Current smoker	7(8.2%)	88(21.4%)	
Missing	0	4	
Family history PD 1st degree (%)*			
Yes	5(8.6%)	24 (6.0%)	0.629

*Only available for follow-up 3 in EPIC-NL, % calculated based on these participants.
PD, Parkinson's Disease; SD, standard deviation; GP, general practitioner.

Table S10. Agreement different sources of PD information and PD status, among cases validated by the GP in the cohorts.

Verified Parkinson status AMIGO		
Ascertainment source	Parkinson (n=18)	No Parkinson (n=208)
Self-reported Parkinson	15 (83.3%)	6 (2.9%)
Self-reported Parkinson medication	15 (83.3%)	11 (5.3%)
PD on death certificate	1 (5.6%)	0 (0%)
EMR registry	16 (88.9%)	98 (47.1%)
EMR ICPC diagnose	14 (77.8%)	3 (1.4%)
EMR ATC medication	14 (77.8%)	97 (46.6%)
Tannerscore 2011 ≥ 5	10 (55.6%)	98 (47.1%)
Tannerscore baseline (mean, SD)	5.00(3.18)	3.52(2.47)
Verified Parkinson status EPIC-NL		
	Parkinson (n=67)	No Parkinson (n=208)
Self-reported Parkinson	43 (64.2%)	29 (13.9%)
Self-reported Parkinson medication	35 (52.2%)	12 (5.8%)
PD on death certificate	9 (13.4%)	1 (0.5%)
HDR registry	23 (34.3%)	19 (9.1%)
Tannerscore 2011 ≥ 5	31 (46.3%)	85 (40.9%)
Tannerscore baseline (mean, SD)	5.85(2.33)	3.86(2.07)
Verified Parkinson status COMBINED		
	Parkinson(n=85)	No Parkinson (n=416)
Self-reported Parkinson	58 (68.2%)	35 (8.4%)
Self-reported Parkinson medication	50 (58.8%)	23 (5.5%)
PD on death certificate	10 (11.8%)	1 (0.2%)
Tannerscore 2011 ≥ 5	41 (48.2%)	183 (44.0%)
Tannerscore baseline (mean, SD)	5.59(2.62)	3.69(2.29)

PD, Parkinson's Disease; EMR, electronic medical records; HDR, hospital discharge registry; SD, standard deviation.

Table S11. Baseline characteristics likelihood 3 versus likelihood 0-2 in the combined cohort.

	likelihood 0-2	**likelihood 3**	**p-value**
Number of participants (%)	54479	346	
Age at baseline			
Mean (SD)	49.55(11.29)	56.71(8.14)	<0.001
Sex (%)			
Male	16720(30.7%)	98(28.3%)	0.372
Female	37759(69.3%)	248(71.7%)	
Education (%)			
Low	28533(52.7%)	198(57.6%)	0.194
Medium	11966(22.1%)	67(19.5%)	
High	13662(25.2%)	79(23.0%)	
Missing	318	2	

	likelihood 0-2	likelihood 3	p-value
Smoking status at baseline (%)			
Never smoker	21828(40.2%)	147(42.9%)	0.002
Past smoker	18046(33.2%)	133(38.8%)	
Current smoker	14421(26.6%)	63(18.4%)	
Missing	184	3	
Family history PD 1st degree (%)*			
Yes	1111(3.9%)	22 (8.9%)	<0.001

*Only available for follow-up 3 in EPIC-NL, % calculated based on these participants. PD, Parkinson's Disease; SD, standard deviation.

Table S12. Baseline characteristics of likelihood 3 versus likelihood 0-2 for EPIC-NL and AMIGO.

	EPIC-NL			AMIGO		
	Likelihood 0-2	Likelihood 3	p-value	Likelihood 0-2	Likelihood 3	p-value
Number of participants (%)	39835	176		14659	170	
Age at baseline						
Mean (SD)	49.16 (11.90)	58.82 (7.48)	<0.001	50.61(9.37)	54.52(8.23)	<0.001
Sex (%)						
Male	10228(25.7%)	32(18.2%)	0.029	6495 (44.3%)	66(38.8%)	0.176
Female	29607(74.3%)	144(81.8%)		8164(55.7%)	104(61.2%)	
Education (%)						
Low	24075(60.9%)	123(70.7%)	0.029	4462(30.5%)	75(44.1%)	0.001
Medium	7384(18.7%)	23(13.2%)		4583(31.3%)	44(25.9%)	
High	8067(20.4%)	28(16.1%)		5605(38.2%)	51(30.0%)	
Missing	309	2		9	0	
Smoking status at baseline (%)						
Never smoker	15157(38.2%)	86(49.4%)	<0.001	6679(45.6%)	61(36.1%)	0.013
Past smoker	12377(31.2%)	63(36.2%)		5674(38.8%)	70(41.4%)	
Current smoker	12139(30.6%)	25(14.4%)		2284(15.6%)	38(22.5%)	
Missing	162	2		22	1	
Family history PD 1st degree (%)*						
Yes	628(4.5%)	10(12.8%)	0.001	484(3.3%)	12(7.1%)	0.013

*Only available for follow-up 3 in EPIC-NL, % calculated based on these participants.
PD, Parkinson's Disease; SD, standard deviation.

Table S13. Crude and adjusted logistic regression analysis of likelihood 3 compared to likelihood 0-2 for the risk factors smoking (baseline), 1st degree family history of PD, and sex in the cohorts.

AMIGO			
	Odds Ratio [95% CI]; crude*	Odds Ratio [95% CI]; adjusted**	Odds Ratio [95% CI]; adjusted***
Smoking at baseline			
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	1.35 [0.96-1.91]	1.10[0.77-1.56]	1.12[0.79-1.59]
Current smokers	1.82[1.20-2.73]	1.69[1.11-2.54]	1.80[1.18-2.69]
1st degree family history of PD			
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.22[1.16-3.86]	1.90[0.99-3.31]	1.92[1.00-3.34]
Sex			
Female	1.0[Ref]	1.0[Ref]	1.0[Ref]
Male	0.81[0.59-1.10]	0.79[0.57-1.07]	0.76[0.55-1.03]
Age			
Age (continuous)	1.05[1.03-1.07]	1.05[1.03-1.07]	1.05[1.03-1.07]
EPIC-NL			
Smoking at baseline			
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	0.90[0.65-1.25]	0.87 [0.62-1.21]	0.87[0.62-1.21]
Current smokers	0.37[0.23-0.56]	0.49[0.30-0.76]	0.49[0.30-0.76]
1st degree family history of PD			
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	3.10[1.49-5.77]	3.09[1.49-5.77]	3.10[1.49-5.79]
Sex (EPIC-MORGEN only)			
Female	1.0[Ref]	1.0[Ref]	1.0[Ref]
Male	1.65[0.95-2.89]	1.56[0.90-2.76]	1.55[0.90-2.72]
Age			
Age (continuous)	1.11[1.09-1.13]	1.11[1.09-1.14]	1.11[1.09-1.14]
COMBINED			
Smoking at baseline			
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	1.09[0.86-1.38]	0.94[0.74-1.19]	0.94[0.74-1.19]
Current smokers	0.77[0.57-1.03]	0.88[0.65-1.18]	0.89[0.66-1.20]
1st degree family history of PD			
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.55[1.59-3.88]	2.29[1.43-3.49]	2.29[1.43-3.49]
Sex			
Female	1.0[Ref]	1.0[Ref]	1.0[Ref]
Male	0.72[0.56-0.91]	0.86[0.67-1.11]	0.85[0.66-1.08]
Age			
Age (continuous)	1.08[1.06-1.09]	1.08[1.06-1.09]	1.08[1.06-1.09]

Likelihood 3 compared to likelihood 0-2. *Adjusted for cohort (only for combined dataset). **Adjusted for age at baseline, baseline education level, sex, cohort(only for combined dataset). *** Adjusted for age at baseline, sex, cohort (only for combined dataset). Ref, reference; PD, Parkinson Disease; CI, Confidence Interval.

Table S14. Crude and adjusted logistic regression analysis of likelihood 3 compared to likelihood 0-1 and likelihood 3 compared to likelihood 0 for the risk factors smoking (baseline), 1st degree family history of PD, and sex in the cohorts.

AMIGO				
	Likelihood 0-1		Likelihood 0	
	Odds Ratio [95% CI]; crude*	Odds Ratio [95% CI]; adjusted**	Odds Ratio [95% CI]; crude*	Odds Ratio [95% CI]; adjusted**
Smoking at baseline				
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	1.36[0.96-1.92]	1.10[0.78-1.57]	1.42[1.00-2.01]	1.13[0.80-1.62]
Current smokers	1.85[1.22-2.77]	1.71[1.12-2.57]	2.02[1.33-3.02]	1.84[1.21-2.77]
1st degree family history of PD				
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.25[1.18-3.90]	1.92[1.00-3.33]	2.31[1.21-4.01]	1.95[1.02-3.41]
Sex				
Female	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Male	0.81[0.59-1.10]	0.79[0.58-1.08]	0.81[0.59-1.10]	0.79[0.58-1.08]
EPIC-NL				
Smoking at baseline				
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	0.91[0.65-1.25]	0.87[0.62-1.21]	0.91[0.66-1.26]	0.88[0.63-1.22]
Current smokers	0.37[0.23-0.56]	0.49[0.30-0.76]	0.36[0.23-0.55]	0.48[0.30-0.74]
1st degree family history of PD				
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	3.10[1.49-5.78]	3.10[1.49-5.79]	3.15[1.51-5.87]	3.16[1.51-5.93]
Sex (EPIC-MORGEN only)				
Female	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Male	1.64[0.95-2.88]	1.55[0.89-2.75]	1.63[0.94-2.85]	1.54[0.88-2.72]
COMBINED				
Smoking at baseline				
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	1.09[0.86-1.38]	0.94[0.74-1.20]	1.12[0.89-1.42]	0.96[0.76-1.22]
Current smokers	0.77[0.57-1.04]	0.88[0.65-1.18]	0.79[0.58-1.06]	0.89[0.66-1.20]
1st degree family history of PD				
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.57[1.60-3.91]	2.30[1.43-3.50]	2.63[1.64-4.00]	2.32[1.45-3.55]
Sex				
Female	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Male	0.72[0.56-0.91]	0.86[0.67-1.11]	0.70[0.55-0.89]	0.85[0.66-1.09]

Likelihood 3 compared to likelihood 0-1 and likelihood 0. *Adjusted for cohort. **Adjusted for age at baseline, baseline education level, sex, cohort. Ref, reference; PD, Parkinson Disease; CI, Confidence Interval.

Table S15. Crude and adjusted logistic regression analysis of confirmed cases by GP for the risk factors smoking (baseline), 1st degree family history of PD, and sex in the cohorts.

COMBINED		
	Odds Ratio [95% CI]; crude*	Odds Ratio [95% CI]; adjusted**
Smoking at baseline		
Never smokers	1.0[Ref]	1.0[Ref]
Past smokers	0.70[0.44-1.11]	0.63[0.39-1.00]
Current smokers	0.20[0.08-0.42]	0.27[0.11-0.56]
1st degree family history of PD		
No first degree family history PD	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.27[0.79-5.19]	2.11[0.73-4.84]
Sex		
Female	1.0[Ref]	1.0[Ref]
Male	0.75[0.44-1.21]	1.34[0.76-2.29]

*Adjusted for cohort. **Adjusted for age at baseline, baseline education level, gender, cohort. Ref, reference; PD, Parkinson Disease; CI, Confidence Interval; GP general practitioner Reference group= likelihood 0.

Table S16. Likelihood 3 with and without medication diagnosis in the electronic medical registry compared to likelihood 0-2 for the risk factors smoking (baseline), 1st degree family history of PD, and sex within AMIGO.

	Odds Ratio [95% CI]; EMR with medication	**Odds Ratio [95% CI]; EMR without medication**
Smoking at baseline		
Never smokers	1.0[Ref]	1.0[Ref]
Past smokers	1.35 [0.96-1.91]	1.27[0.73-2.23]
Current smokers	1.82[1.20-2.73]	2.06[1.09-3.83]
1st degree family history of PD		
No first degree family history PD	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.22[1.16-3.86]	2.34[0.82-5.30]
Sex		
Female	1.0[Ref]	1.0[Ref]
Male	0.81[0.59-1.10]	1.65[1.02-2.71]

* Controls were all participants with likelihood score 0-2.
LH, likelihood; Ref, reference; PD, Parkinson Disease; CI, Confidence Interval.

Table S17. Logistic regression analysis of likelihood 3 compared to likelihood 0 for the risk factors smoking (baseline), 1st degree family history of PD in the cohorts stratified by sex and age.

AMIGO				
	female	male	< 60 years*	>=60 years
Smoking at baseline				
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	1.43 [0.92-2.23]	1.40[0.80-2.47]	1.46[0.93-2.30]	1.00[0.58-1.74]
Current smokers	1.93[1.12-3.25]	2.17[1.12-4.10]	2.55[1.57-4.13]	1.13[0.47-2.45]
1st degree family history of PD				
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	3.09[1.44-5.89]	1.33[0.32-3.63]	2.82[1.25-5.50]	1.37[0.41-3.40]
EPIC-NL				
Smoking at baseline				
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	0.91[0.63-1.31]	1.01[0.46-2.20]	0.71[0.44-1.12]	1.30[0.81-2.08]
Current smokers	0.42[0.25-0.67]	0.24[0.07-0.69]	0.35[0.19-0.61]	0.63[0.28-1.26]
1st degree family history of PD				
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	3.21[1.47-6.21]	2.39[0.13-12.49]	2.53[0.97-5.48]	5.41[1.52-15.25]
COMBINED				
Smoking at baseline				
Never smokers	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
Past smokers	1.09[0.83-1.44]	1.27[0.81-2.00]	1.04[0.75-1.42]	1.14 [0.80-1.64]
Current smokers	0.75[0.52-1.06]	1.02[0.57-1.78]	0.95[0.66-1.37]	0.79[0.45-1.33]
1st degree family history of PD				
No first degree family history PD	1.0[Ref]	1.0[Ref]	1.0[Ref]	1.0[Ref]
First degree family history of PD	2.92[1.72-4.67]	1.48[0.45-3.60]	2.48[1.37-4.16]	2.23[0.98-4.42]

Figure S1. Venn diagram of different sources of PD information in AMIGO.

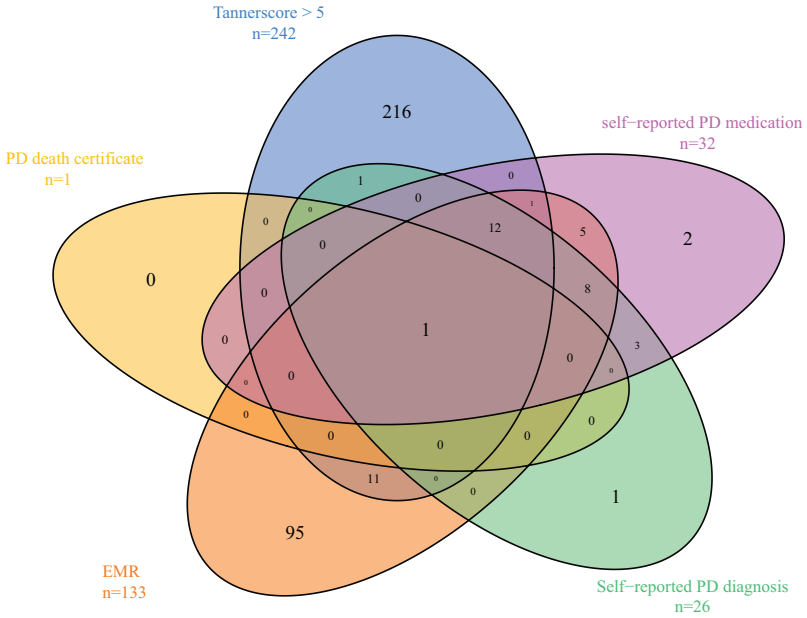
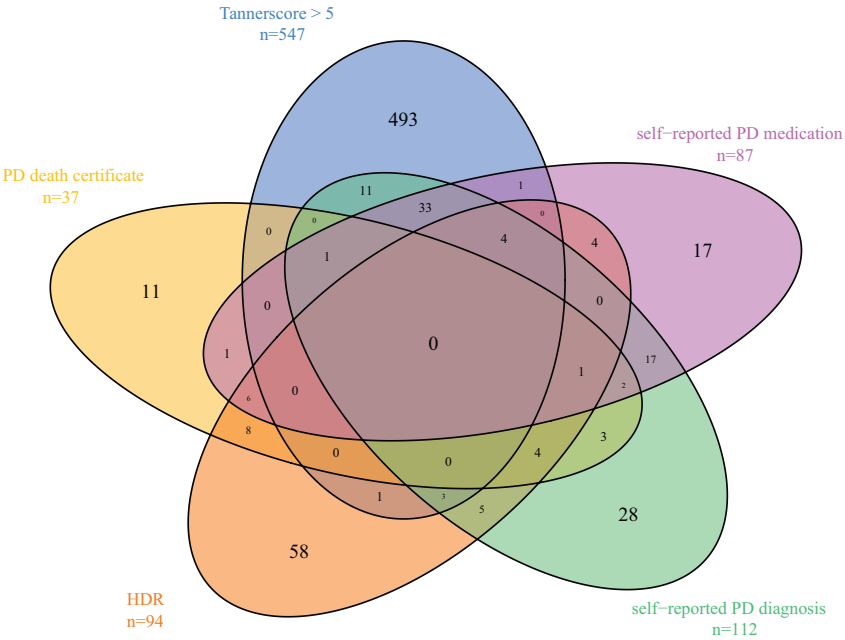


Figure S2. Venn diagram of different sources of PD information in EPIC-NL.



REFERENCES

1. de Rijk MC, Launer LJ BK. Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Neurology. 2000;54: S21–23.
2. Jankovic J. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry. 2008;79: 368–376. doi:10.1136/jnnp.2007.131045
3. Obeso J a, Rodriguez-Oroz MC, Goetz CG, Marin C, Kordower JH, Rodriguez M, et al. Missing pieces in the Parkinson's disease puzzle. Nat Med. 2010;16: 653–661. doi:10.1038/nm.2165
4. Samii A, Nutt JG, Ransom BR. Parkinson's disease. Lancet. 2004;363: 1783–1793. doi:10.1016/S0140-6736(04)16305-8
5. Ritz B, Ascherio A, Checkoway H, Marder KS, Nelson LM, Rocca WA, et al. Pooled Analysis of Tobacco Use and Risk of Parkinson Disease. Arch Neurol. 2007;64: 990. doi:10.1001/archneur.64.7.990
6. Taylor KSM, Cook JA, Counsell CE. Heterogeneity in male to female risk for Parkinson's disease. J Neurol Neurosurg & Psychiatry. 2007;78: 905–906. doi:10.1136/jnnp.2006.104695
7. Wirdefeldt K, Adami H-O, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. Eur J Epidemiol. 2011;26: 1–58. doi:10.1007/s10654-011-9581-6
8. Brouwer M, Koeman T, van den Brandt PA, Kromhout H, Schouten LJ, Peters S, et al. Occupational exposures and Parkinson's disease mortality in a prospective Dutch cohort. Occup Environ Med. 2015;72: 448–455. doi:10.1136/oemed-2014-102209
9. Tysnes O-B, Storstein A. Epidemiology of Parkinson's disease. J Neural Transm. 2017;124: 901–905. doi:10.1007/s00702-017-1686-y
10. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. Lancet Neurol. 2006;5: 525–535. doi:10.1016/S1474-4422(06)70471-9
11. Rascol O, Lozano A, Stern M, Poewe W. Milestones in Parkinson's disease therapeutics. Mov Disord. 2011;26: 1072–82. doi:10.1002/mds.23714
12. Ton TGN, Biggs M Lou, Comer D, Curtis L, Hu S-C, Thacker EL, et al. Enhancing case ascertainment of Parkinson's disease using Medicare claims data in a population-based cohort: the Cardiovascular Health Study. Pharmacoepidemiol Drug Saf. 2014;23: 119–127. doi:10.1002/pds.3552
13. Tanner CM, Gilley DW GC. A brief screening questionnaire for parkinsonism. Ann Neurol. 1990;28: 267–8.
14. Noyce AJ, Bestwick JP, Silveira-Moriyama L, Hawkes CH, Giovannoni G, Lees AJ, et al. Meta-analysis of early nonmotor features and risk factors for Parkinson disease. Ann Neurol. 2012;72: 893–901. doi:10.1002/ana.23687
15. Beulens JWJ, Monninkhof EM, Verschuren WMM, Schouw YT v. d., Smit J, Ocke MC, et al. Cohort Profile: The EPIC-NL study. Int J Epidemiol. 2010;39: 1170–1178. doi:10.1093/ije/dyp217
16. Slottje P, Yzermans CJ, Korevaar JC, Hooiveld M, Vermeulen RCH. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in the Netherlands. BMJ Open. 2014;4: e005858. doi:10.1136/bmjopen-2014-005858
17. Rocca W a, Maraganore DM, McDonnell SK, Schaid DJ. Validation of a Telephone Questionnaire for Parkinson ' s Disease. J Clin Epidemiol. 1998;51: 517–523.
18. Dahodwala N. Screening Questionnaires for Parkinsonism: A Systematic Review. Park Relat Disord. 2012;29: 997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted
19. Höglinger GU, Rissling I, Metz A, Ries V, Heinermann A, Prinz H, et al. Enhancing recognition of early Parkinsonism in the community. Mov Disord. 2004;19: 505–512. doi:10.1002/mds.20033
20. Barbosa MT, Caramelli P, Maia DP, Cunningham MCQ, Guerra HL, Lima-Costa MF, et al. Parkinsonism and Parkinson's disease in the elderly: A community-based survey in Brazil (the Bambuí Study). Mov Disord. 2006;21: 800–808. doi:10.1002/mds.20806

21. Taylor KSM, Counsell CE, Harris CE, Gordon JC. Screening for undiagnosed parkinsonism in people aged 65 years and over in the community. 2006;12: 79–85. doi:10.1016/j.parkreldis.2005.08.005

22. Gallo V, Brayne C, Forsgren L, Barker RA, Petersson J, Hansson O, et al. Parkinson's Disease Case Ascertainment in the EPIC Cohort: The NeuroEPIC4PD Study. Neurodegener Dis. 2015;15: 331–338. doi:10.1159/000381857

23. Rijk M De, Breteler M, Graveland G. Prevalence of Parkinson's disease in the elderly The Rotterdam Study. Neurology. 1995.

24. de Lau LML, Giesbergen PCLM, de Rijk MC, Hofman a, Koudstaal PJ, Breteler MMB. Incidence of parkinsonism and Parkinson disease in a general population: the Rotterdam Study. Neurology. 2004;63: 1240–1244. doi:10.1212/01.WNL.0000140706.52798.BE

25. Struijk EA, May AM, Beulens JWJ, Van Gils CH, Monninkhof EM, Van Der Schouw YT, et al. Mortality and cancer incidence in the EPIC-NL cohort: Impact of the healthy volunteer effect. Eur J Public Health. 2015;25: 144–149. doi:10.1093/eurpub/cku045

26. Wooten GF. Are men at greater risk for Parkinson's disease than women? J Neurol Neurosurg Psychiatry. 2004;75: 637–639. doi:10.1136/jnnp.2003.020982

27. Jain S, Himali J, Beiser A, Ton TGN, Kelly-Hayes M, Biggs M Lou, et al. Validation of Secondary Data Sources to Identify Parkinson Disease Against Clinical Diagnostic Criteria. Am J Epidemiol. 2014;181: 185–190. doi:10.1093/aje/kwu326

28. Ton TG, Jain S, Boudreau R, Thacker EL, Strotmeyer ES, Newman AB, et al. Post Hoc parkinson's disease: Identifying an uncommon disease in the cardiovascular health study. Neuroepidemiology. 2010;35: 241–249. doi:10.1159/000319895

29. Wirdefeldt K, Gatz M, Bakaysa SL, Fiske A, Flensburg M, Petzinger GM, et al. Complete ascertainment of Parkinson disease in the Swedish Twin Registry. Neurobiol Aging. 2008;29: 1765–1773. doi:10.1016/j.neurobiolaging.2007.04.009

30. Swarztrauber K, Anau J, Peters D. Identifying and distinguishing cases of parkinsonism and Parkinson's disease using ICD-9 CM codes and pharmacy data. Mov Disord. 2005;20: 964–970. doi:10.1002/mds.20479

31. Karin Wirdefeldt, Hans-Olov Adami, Philip Cole DT and JM. Epidemiology and etiology of Parkinson's disease: a review of the evidence. Eur J Epidemiol. 2011;26: 385–393. doi:10.1007/sl0654-01

32. Sääksjärvi K, Knekt P, Rissanen H, Laaksonen M a, Reunanen a, Männistö S. Prospective study of coffee consumption and risk of Parkinson's disease. Eur J Clin Nutr. 2008;62: 908–915. doi:10.1038/sj.ejcn.1602788

Chapter 5

Associations of Maternal Cell-Phone Use During Pregnancy With Pregnancy Duration and Fetal Growth in 4 Birth Cohorts

Ermioni Tsarna, Marije Reedijk, Laura Ellen Birks, Mònica Guxens, Ferran Ballester, Mina Ha, Ana Jiménez-Zabala, Leeka Kheifets, Aitana Lertxundi, Hyung-ryul Lim, Jorn Olsen, Llúcia González Safont, Madhuri Sudan, Elisabeth Cardis, Martine Vrijheid, Tanja Vrijkotte, Anke Huss*, Roel Vermeulen*

*Equal contribution

Published: American Journal of Epidemiology, Volume 188, Issue 7, July 2019, Pages 1270–1280

ABSTRACT

Results from studies evaluating potential effects of prenatal exposure to radio-frequency electromagnetic fields from cell phones on birth outcomes have been inconsistent. Using data on 55,507 pregnant women and their children from Denmark (1996–2002), the Netherlands (2003–2004), Spain (2003–2008), and South Korea (2006–2011), we explored whether maternal cell-phone use was associated with pregnancy duration and fetal growth. On the basis of self-reported number of cell-phone calls per day, exposure was grouped as none, low (referent), intermediate, or high. We examined pregnancy duration (gestational age at birth, preterm/postterm birth), fetal growth (birth weight ratio, small/large size for gestational age), and birth weight variables (birth weight, low/high birth weight) and meta-analyzed cohort-specific estimates. The intermediate exposure group had a higher risk of giving birth at a lower gestational age (hazard ratio = 1.04, 95% confidence interval: 1.01, 1.07), and exposure-response relationships were found for shorter pregnancy duration ($P < 0.001$) and preterm birth ($P = 0.003$). We observed no association with fetal growth or birth weight. Maternal cell-phone use during pregnancy may be associated with shorter pregnancy duration and increased risk of preterm birth, but these results should be interpreted with caution, since they may reflect stress during pregnancy or other residual confounding rather than a direct effect of cell-phone exposure.

Abbreviations

HR, hazard ratio; RF-EMF, radiofrequency electromagnetic fields; 3G, third generation, 2G, second generation, SGA, small for gestational age; DNBC, Danish National Birth Cohort; ABCD, Amsterdam Born Children and their Development Study; INMA, Spanish Environment and Childhood Project; MOCEH, Korean Mothers and Children's Environment Health Study; n, sample size; SD, standard deviation; LGA, large for gestational age; BMI, body mass index, PH, proportional hazards; OR, odds ratio; HR, hazard ratio; TSH, thyroid stimulating hormone; CRH, corticotrophin releasing hormone;

INTRODUCTION

Cell phone use has rapidly increased during the last decades[1]. Cell phones generate radio-frequency electromagnetic fields (RF-EMF), resulting in local exposure of the human body to RF-EMF. Concerns have been raised regarding the potential effects of RF-EMF exposure on human health. Short-term exposure in adults is regarded as safe; however, potential health effects of long-term exposure, and exposure on fetuses and children are not well studied[2,3].

The interpretation of results from epidemiologic studies evaluating health effects of RF-EMF exposure from cell phone use is challenging. Because of the technical evolution of mobile communication systems, similar levels of cell phone use do not necessarily induce similar levels of RF-EMF exposure. User's exposure to RF-EMF from third generation devices (3G) is much lower than from second generation (2G) devices[4]. More recent technologies (4G/5G) also differ from 3G devices, though less, in terms of user's exposure.

Exposure to RF-EMF during pregnancy could affect the growth and development of the fetus and the pregnancy duration, either directly due to radiation of the fetus and the placenta or indirectly as a result of altered maternal physiology. Some animal studies have shown an association on steroidogenesis and lower birth weight of the offspring[5,6]. However, not all animal studies support an association with adverse birth outcomes[7,8]. During cell-phone calling and texting, abdominal exposure is low and modeling studies estimate that exposure levels of the human fetus are very low[9-12], although an experimental study in humans has shown that abdominal RF-EMF exposure may affect the placenta function[13]. In addition, an association between RF-EMF exposure and thyroid dysfunction has been indicated in animal studies[14,15].

Previous epidemiologic studies have given inconsistent results. In a cohort study from Turkey (N=500), Col-Araz retrospectively assessed cell-phone use and reported shorter pregnancy duration and increased risk for preterm birth[16]. In a cohort study from Iran (N=1,200), Mortazavi et al found no association with birth weight[17]. In a much larger sample from Norway (N=100,231), Baste et al. found no association between cell-phone use and low birth weight, preterm birth, or small-for-gestational-age (SGA) birth[18].

Taking into account the ubiquity of cell-phone use, an effect on birth outcomes, even small, may have considerable public health impact. Our aim of this study was to explore the possible association of maternal cell-phone use with pregnancy duration and fetal growth in 4 birth cohorts from Denmark, the Netherlands, Spain, and South Korea.

METHODS

Study population

Our study was conducted within 4 population-based birth cohorts, participating in the Generalized EMF Research using Novel Methods (GERoNiMO) Project[19] - namely, the Danish National Birth Cohort (DNBC)[20,21], the Amsterdam Born Children and their Development Study (ABCD)[22], the Spanish Environment and Childhood Project (INMA)[23], and the Korean Mothers and Children's Environment Health Study (MOCEH)[24]. In total, 113,319 pregnant women were enrolled from 1996 to 2011 in these cohorts. All participants gave written informed consent, and each cohort was ethically approved from local research ethical committees. Our inclusion criteria were the availability of information on frequency of cell-phone calls (both incoming and outgoing) during pregnancy, child's birth weight, and gestational age at birth. Mother-child pairs were excluded in case of multiple pregnancy, if a spontaneous abortion occurred (gestational age at birth<20 weeks, n=39), or records of birth outcomes were implausible (i.e. if gestational age at birth was ≥47 weeks (n=27) and/or the birth weight was more than four standard deviations (SD) away from the mean for gestational age, based on birth-weight reference curves (n=126)). National sex-specific birth-weight reference curves relevant to the study period were available for the Netherlands (ABCD)[25], Spain (INMA)[26], and Korea (MOCEH)[27]. In DNBC, the Norwegian reference curves were used, because of unavailability of national Danish curves developed with similar methodology and relevant to the study period[28]. In total 55,507 mother-child pairs met our inclusion criteria (Table 1).

Maternal cell phone use during pregnancy

The mothers from DNBC and ABCD reported their frequency of cell-phone calls during pregnancy 7 years postnatally. In INMA and MOCEH, similar questionnaires were given to the mothers during pregnancy. To be consistent with previous analyses within these cohorts[29], we classified exposure in 4 categories (none, low, intermediate, high), based on available information regarding daily frequency of cell-phone calls during pregnancy (Table 2). During the enrollment period of DNBC 2G devices were used; in the more recent cohorts, 3G and 2G devices were used alongside each other. Thus, RF-EMF exposure from similar cell-phone use should have been higher in DNBC, on average, and lower in the more recent cohorts.

Table 1. Availability of Data on Exposure and Outcomes and the Study Population for 4 Cohorts Included in an Analysis of Maternal Cell-Phone Use During Pregnancy and Birth Outcomes, 1996–2011

Cohort	Location	Enrolment		Cell phone use during pregnancy		Pregnancy Duration and Fetal Growth Outcomes	Study population (% of N enrolled)
		Time period	No, of Pairs enrolled[a]	Time of data collection	N (% of N enrolled)	N (% of N enrolled)	
DNBC	Denmark	1996-2002	101,032	7 years postnatal	50,040 [b](49.5)	54,498 (53.9)[a]	49,668 (49.2)
ABCD	The Netherlands	2003-2004	8,266	7 years postnatal	2,611 (31.6)	7,812 (94.5)	2,597 (31.4)
INMA	Spain	2003-2008	2,270	pregnancy	1,993 (87.8)	1,975 (87.0)	1,934 (85.2)
MOCEH	South Korea	2006-2011	1,751	pregnancy	1,435 (82.0)	1,352 (77.2)[c]	1,308 (74.7)
Total N			113,319		56,079 (49.5)	65,637 (57.9)	55,507 (49.0)

Abbreviations: ABCD, Amsterdam Born Children and Their Development Study; DNBC, Danish National Birth Cohort; INMA, Spanish Environment and Childhood Project; MOCEH, Korean Mothers and Children's Environment Health Study.
[a] Number of mother-child pairs.
[b] Out of 54,908 offspring whose mothers responded to the age 7 years questionnaire.
[c] Out of 1,481 offspring whose mothers responded to the cell-phone use questionnaire.

Table 2. Classification of Cell-Phone Exposure in 4 Cohorts Included in an Analysis of Maternal Cell-Phone Use During Pregnancy and Birth Outcomes, 1996–2011

Exposure's classification[a]	**DNBC N[b](%)**	**ABCD N(%)**	**INMA N(%)**	**MOCEH N(%)**	**Total N(%)**
None	30,185 (60.8)	180 (6.9)	53 (2.7)	15 (1.2)	30,433 (54.8)
Low	10,860 (21.9)	1,125 (43.3)	703 (36.4)	242 (18.5)	12,930 (23.3)
Intermediate	6,172 (12.4)	703 (27.1)	753 (38.9)	642 (49.1)	8,270 (14.9)
High	2,451 (4.9)	589 (22.7)	425 (22.0)	409 (31.3)	3,874 (7.0)
TOTAL	49,668	2,597	1,934	1,308	55,507

Abbreviations: ABCD, Amsterdam Born Children and Their Development Study; DNBC, Danish National Birth Cohort; INMA, Spanish Environment and Childhood Project; MOCEH, Korean Mothers and Children's Environment Health Study.
[a] In the DNBC, ABCD, and INMA cohorts, no exposure corresponded to no cell-phone use, low exposure to ≤1 calls/day, intermediate exposure to 2–3 calls/day, and high exposure to ≥4 calls/day. In the MOCEH cohort, no exposure corresponded to no cell-phone use, low exposure to ≤2 calls/day, intermediate exposure to 3–5 calls/day, and high exposure to ≥6 calls/day.
[b] Number of mother-child pairs.

Pregnancy duration and fetal growth outcome

We defined all outcomes of this study a priori. We examined pregnancy duration, using gestational age at birth, preterm (≤ 36 completed weeks) and postterm birth (>42 completed weeks); fetal growth, using birth weight ratio, SGA birth and large-for-gestational-age (LGA) birth; and birth weight, low birth weight (≤ 2.499 g), and high birth weight (≥ 4.000 g), which reflect both pregnancy duration and fetal growth. Birth weight ratio was defined as the observed birth weight divided by the median birth weight from a national birth weight reference curve[30]; SGA birth as birth weight below the 10th percentile, and LGA birth above the 90th percentile. In DNBC, gestational age at birth was reported by midwifes based on the woman's last menstrual cycle and ultrasound examinations. In INMA, the date of last menstrual cycle was used, if this was consistent with the ultrasound-based estimate (≤7days' difference); otherwise the ultrasound estimate was used. Women for whom this difference exceeded 3 weeks were removed from the study. In ABCD and MOCEH, gestational age at birth was defined on the basis of ultrasound examinations during pregnancy; if this information was not available, gestational age at birth was calculated based on last menstrual period.

Covariate data

We preselected the following covariates for the adjusted statistical models: maternal age at child's birth (a natural spline term with 3 degrees of freedom), parity, active and passive smoking during pregnancy, alcohol consumption during pregnancy, prepregnancy body mass index (weight (kg)/height (m)^2; a natural spline term with knots at cutoff values between underweight, normal, and overweight as appropriate for Caucasian and Asian populations)[31], height, educational level, socioeconomic position, and marital status. In addition, geographical region was a covariate for the analysis within the multicentre cohorts (INMA and MOCEH) and maternal country of birth (European/non-European) for the analysis within DNBC, ABCD and INMA cohorts, where this was heterogeneous. Data regarding the aforementioned variables were self-reported in questionnaires or telephone interviews during pregnancy or after birth. Definitions of covariates per cohort are provided in Web Tables 1-4.

Statistical analysis

Multiple imputation by chained equations was used for missing values in the covariates of the adjusted statistical models. It was performed in the study population and separately for each cohort. All covariates -apart from geographical region, and all study outcomes were used as predictors, and 20 complete datasets were obtained[32]. In each cohort, maternal characteristics and pregnancy duration and fetal growth outcomes were characterized by exposure group, using mean values and proportions as appropriate. Modified Wald test statistic[33], Chi-square, and Fischer exact tests were performed to detect any difference between the exposure groups (Web Tables 1-4).

For the analysis of birth weight and birth weight ratio, we used multiple linear regression models. To achieve normality of residuals, we excluded preterm neonates from the analysis of birth weight. Gestational age at birth was treated as time from conception to birth, and Cox proportional hazards (PH) models were used[34]. To meet the assumption of proportional hazards, we used parity status, active and passive smoking, and alcohol consumption as stratifying variables. For the analysis of low and high birth weight, preterm and postterm birth, and SGA and LGA birth, we used logistic regression models. In all statistical models, the low exposure group was the reference, because of the very low proportion of women reporting no use of cell-phone during pregnancy in ABCD, INMA, and MOCEH (Table 2).

The calculated unadjusted and adjusted cohort-specific estimates were meta-analyzed those using random-effects models. INMA was excluded from the meta-analysis of odds ratios of the unexposed group for postterm birth, because there were no cases in this group (Web Table 3). Similarly, MOCEH was excluded from the meta-analysis of odds ratio's of the unexposed group for postterm birth and low birth weight (Web Table 4). We refitted the adjusted statistical models described above with a continuous exposure variable and meta-analyzed the obtained effect estimates with random-effects model. The corresponding *P*-value is the reported statistical significance of the linear trend.

We performed the following sensitivity analysis: 1) complete-case analysis, to assess the influence of multiple imputation; 2) analysis with binary exposure (none/low vs. intermediate/high), to achieve maximum statistical power while including the unexposed and highly exposed mothers in the comparisons; 3) analysis of low birth weight restricted to nonpreterm neonates, to assess whether the results of primary analysis were driven by preterm births; 4) analysis of birth weight ratio, SGA, and LGA in DNBC using the observed birth weight percentiles per gestational age, to assess the impact of using the Norwegian reference curves in our primary analysis; 5) meta-analysis of results excluding one cohort at a time, to assess the influence of each cohort on our pooled estimates; and 6) meta-analysis of cohorts with retrospective exposure assessment (DNBC and ABCD) versus prospective exposure assessment (INMA and MOCEH), to assess the effect of recall error.

All analyses were performed using R statistical software (version 3.4.0)[35], and the following packages: "tableone"[36], "mice"[37], "miceadds"[38], "splines"[35], "survival"[30,40], and "metafor" [41].

RESULTS

Descriptive statistics

In our study population of 55,507 mother-child pairs, mean birth weight was 3,578 (standard deviation (SD), 547) g; 1,448 (2.61%) children were born with low birth weight and 12,188 (21.96%) with high birth weight (Table 3). The average gestational age at birth was 39.98 (SD, 1.67) weeks; 2,271 (4.09%)

children were born preterm and 3,170 (5.71%) postterm. The distribution of gestational age at birth was left-skewed and indicative of right-censoring, because of the cesarean deliveries and induced vaginal labors. The incidence of postterm birth varied between 6% in DNBC and 0.6% in MOCEH. Regarding fetal growth, the mean birth weight ratio was 1.01 (SD, 0.13); 3,535 children (6.37%) were born SGA and 8,287 (14.93%) LGA. Incidence of SGA and LGA was closer to the expected 10% in the ABCD, INMA, and MOCEH cohorts.

Table 3. Pregnancy Duration and Fetal Growth Outcomes in an Analysis of Maternal Cell-Phone Use During Pregnancy and Birth Outcomes, 1996–2011

Birth Outcomes	DNBC N[a](%)	ABCD N(%)	INMA N(%)	MOCEH N(%)	Total N(%)
Birth Weight, g[b]	3,602 (547)	3,503 (521)	3,262 (461)	3,261 (441)	3,578 (547)
Low birth weight (≤2,499 g)	1,248 (2.51)	77 (2.96)	89 (4.60)	34 (2.60)	1,448 (2.61)
High birth weight (≥4,000 g)	11,598 (23.35)	424 (16.33)	108 (5.58)	58 (4.43)	12,188 (21.96)
Gestational age at birth, weeks[b]	40.00 (1.67)	39.93 (1.61)	39.83 (1.60)	39.19 (1.60)	39.98 (1.67)
Preterm (≤36 completed weeks)	2,025 (4.08)	117 (4.51)	65 (3.36)	64 (4.89)	2,271 (4.09)
Postterm (>42 completed weeks)	2,982 (6.00)	108 (4.16)	72 (3.72)	8 (0.61)	3,170 (5.71)
Birth Weight Ratio [b,c]	1.01 (0.13)	1.01 (0.12)	1.01 (0.12)	1.01 (0.12)	1.01 (0.13)
SGA birth (<10th percentile)	3,030 (6.10)	206 (7.93)	182 (9.41)	117 (8.94)	3,535 (6.37)
LGA birth (>90th percentile)	7,660 (15.42)	276 (10.63)	205 (10.60)	146 (11.16)	8,287 (14.93)
TOTAL n	49,668	2,597	1,934	1,308	55,507

Abbreviations: ABCD, Amsterdam Born Children and Their Development Study; DNBC, Danish National Birth Cohort; INMA, Spanish Environment and Childhood Project; LGA, large for gestational age; MOCEH, Korean Mothers and Children's Environment Health Study; SD, standard deviation; SGA, small for gestational age.
[a] Number of mother-child pairs.
[b] Values are expressed as mean (standard deviation).
[c] Birth weight ratio was defined as observed birth weight divided by the median birth weight from a national birth-weight reference curve (30).

With respect to maternal cell-phone use during pregnancy, 55% of the mothers were classified in the unexposed group, 23% in the low-exposure group, 15% in the intermediate-exposure group, and 7% in the high-exposure group (Table 2). In the older cohort (DNBC), cell-phone use was less frequent (61% unexposed). In all 4 cohorts, mothers with higher cell-phone use during pregnancy were more often primiparous, were more likely to smoke during pregnancy, and were more likely to be exposed to secondhand smoke (Web Tables 1 - 4). In ABCD, INMA, and MOCEH cohorts, higher maternal cell-phone use was associated with a higher educational level; however, the opposite was seen in the DNBC cohort.

Adjusted associations of maternal cell-phone use with pregnancy duration and fetal growth outcomes

With respect to pregnancy duration, the intermediate-exposure group had a higher risk of giving birth at a lower gestational age compared with the low-exposure group (hazard ratio (HR) = 1.04, 95%

confidence interval (CI): 1.01, 1.07) (Table 4, Web Figure 1). The hazard ratios for the other exposure groups were closer to unity (unexposed: HR = 0.99 (95% CI: 0.97, 1.01); highly exposed: HR = 1.02 (95% CI: 0.98, 1.06)), but a linear trend was observed (P < 0.001). In the analysis of preterm birth, a linear trend was observed (P = 0.003), though none of the odds ratios reached statistical significance (unexposed: odds ratio (OR) = 0.96 (95% CI: 0.86, 1.07); intermediate exposure: OR = 1.12 (95% CI: 0.97, 1.28); highly exposed: OR = 1.28 (95% CI: 0.87, 1.88)) (Table 4, Figure 1). For postterm birth, a significant odds ratio (OR = 0.85, 95% CI: 0.75, 0.97) was observed only for the intermediate-exposure group, but there was no linear trend in the results (P = 0.86) (Table 4, Web Figure 2). No association of maternal cell-phone use with fetal growth was detected in any of the examined outcomes (birth weight ratio, SGA, and LGA) (Table 4, Web Figures 3–5). Regarding birth weight, no association or linear trend was observed within the nonpreterm neonates (Web Figure 6); similarly, the odds of high birth weight did not differ from unity (Web Figure 7). In the analysis of low birth weight, we observed a significant decrease in the odds for the unexposed group (OR = 0.87, 95% CI: 0.76, 1.00) and a linear trend (P = 0.01) (Table 4, Web Figure 8). Note that 65% (n = 947) of the low-birth-weight cases were also born preterm. All cohort-specific unadjusted and adjusted estimates are shown in Web Tables 5 and 6.

Table 4. Results From a Meta-Analysis of the Associations of Maternal Cell-Phone Use During Pregnancy With Pregnancy Duration and Fetal Growth Outcomes, 1996–2011

Birth Outcome[a] and Category of Maternal Cell-Phone Use[b]	**Maternal cell phone use**	**Cases**	**Unadjusted results**	**Adjusted results**	**p-value for trend[c]**
			MD 95%CI	**MD[c] 95%CI**	
Birth Weight in non preterm neonates, g	None		**11.68 (0.81, 22.54)**	-11.15(-53.24,30.94)	0.093
	Low		reference	reference	
	Intermediate		-16.20 (-35.13, 2.73)	-8.17(-21.34,5.00)	
	High		-11.84 (-29.88, 6.20)	-2.56(-19.90,14.78)	
Birth Weight Ratio	None		**0.00 (0.00,0.01)**	0.00(-0.00,0.00)	0.392
	Low		reference	reference	
	Intermediate		-0.00 (-0.01,0.01)	-0.00(-0.00,0.00)	
	High		-0.00 (-0.01, 0.00)	0.00(-0.00,0.00)	
			OR 95%CI	**OR[c] 95%CI**	
Low Birth Weight	None	684	**0.81 (0.71, 0.93)[d]**	**0.87 (0.76, 1.00)[d]**	**0.011**
	Low	373	reference	reference	
	Intermediate	251	1.04 (0.89, 1.23)	0.95(0.81,1.13)	
	High	140	**1.23 (1.01, 1.51)**	1.13(0.92,1.40)	
High Birth Weight	None	7,244	1.02 (0.87, 1.20)	0.92(0.68,1.24)	0.268
	Low	2,739	reference	reference	
	Intermediate	1,543	0.92 (0.82, 1.03)	0.96(0.83,1.10)	
	High	662	0.89 (0.77, 1.04)	0.93(0.78,1.11)	

Birth Outcome[a] and Category of Maternal Cell-Phone Use[b]	Maternal cell phone use	Cases	Unadjusted results	Adjusted results	p-value for trend[c]
			MD 95%CI	**MD[c] 95%CI**	
Preterm birth	None	1,145	**0.90 (0.80, 1.00)**	0.96(0.86,1.07)	**0.003**
	Low	539	reference	reference	
	Intermediate	393	**1.17 (1.02, 1.34)**	1.12(0.97,1.28)	
	High	194	**1.21 (1.02, 1.44)**	1.28(0.87,1.88)	
Postterm birth	None	1,799	0.92 (0.84, 1.00)[e]	0.98 (0.89, 1.07) [e]	0.863
	Low	770	reference	reference	
	Intermediate	400	**0.87 (0.77, 0.98)**	**0.85(0.75,0.97)**	
	High	201	1.06 (0.82, 1.37)	0.98(0.83,1.16)	
SGA Birth	None	1,779	**0.90 (0.82, 0.98)**	0.94(0.86,1.03)	0.872
	Low	877	reference	reference	
	Intermediate	608	1.05 (0.87, 1.26)	1.03(0.88,1.21)	
	High	272	0.95 (0.82, 1.10)	0.94(0.78,1.13)	
LGA Birth	None	4,773	1.01 (0.95, 1.08)	0.98(0.92,1.04)	0.488
	Low	1,916	reference	reference	
	Intermediate	1,112	**0.92 (0.85, 1.00)**	0.97(0.89,1.05)	
	High	490	**0.89 (0.80, 0.99)**	0.93(0.83,1.04)	
			HR 95CI%	**HR[c] 95%CI**	
Gestational age at birth (Hazard Ratio)	None		1.00 (0.98, 1.02)	0.99 (0.97,1.01)	**<0.001**
	Low		reference	reference	
	Intermediate		1.01 (0.96, 1.06)	**1.04 (1.01,1.07)**	
	High		1.01 (0.98, 1.05)	1.02 (0.98,1.06)	

Abbreviations: ABCD, Amsterdam Born Children and Their Development Study; CI, confidence interval; DNBC, Danish National Birth Cohort; HR, hazard ratio; INMA, Spanish Environment and Childhood Project; LGA, large for gestational age; MD, mean difference; MOCEH, Korean Mothers and Children's Environment Health Study; OR, odds ratio; SGA, small for gestational age.

[a]For definitions of birth outcomes, see Table 3.

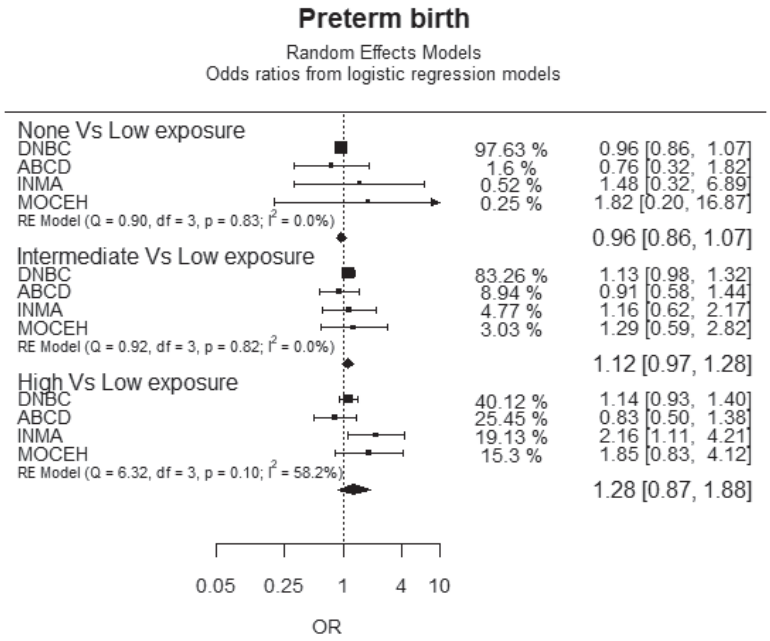
[b] In the DNBC, ABCD, and INMA cohorts, no exposure corresponded to no cell-phone use, low exposure to ≤1 calls/day, intermediate exposure to 2–3 calls/day, and high exposure to ≥4 calls/day. In the MOCEH cohort, no exposure corresponded to no cell-phone use, low exposure to ≤2 calls/day, intermediate exposure to 3–5 calls/day, and high exposure to ≥6 calls/day.

[c] Adjusted for maternal age, parity, active and passive smoking, alcohol consumption, prepregnancy body mass index, educational level, socioeconomic position, marital status, and maternal height.

[d] Excluding MOCEH.

[e]Excluding MOCEH and INMA

Figure 1.



Odds ratios (ORs) from a meta-analysis of the association of maternal cell-phone use during pregnancy with the odds of giving birth preterm. The meta-analysis included data on 55,507 pregnant women and their children from Denmark (1996–2002), the Netherlands (2003–2004), Spain (2003–2008), and South Korea (2006–2011). In the DNBC, ABCD, and INMA cohorts, no exposure corresponded to no cell-phone use, low exposure to ≤1 call/day, intermediate exposure to 2–3 calls/day, and high exposure to ≥4 calls/day. In the MOCEH cohort, no exposure corresponded to no cell-phone use, low exposure to ≤2 calls/day, intermediate exposure to 3–5 calls/day, and high exposure to ≥6 calls/day. Low exposure was the referent group for all cohorts. Squares show individual study estimates; diamonds show pooled estimates. Results were adjusted for maternal age, parity, active and passive smoking, alcohol consumption, prepregnancy body mass index (weight (kg)/height (m)^2), educational level, socioeconomic position, marital status, and maternal height. For no exposure, Q = 0.90 (3 degrees of freedom (df)), P = 0.83, and I^2 = 0.0%; for intermediate exposure, Q = 0.92 (3 df), P = 0.82, and I^2 = 0.0%; and for high exposure, Q = 6.32 (3 df), P = 0.10, and I^2 = 58.2%. Bars, 95% confidence intervals (CIs). ABCD, Amsterdam Born Children and Their Development Study; DNBC, Danish National Birth Cohort; INMA, Spanish Environment and Childhood Project; MOCEH, Korean Mothers and Children's Environment Health Study.

Sensitivity analyses

In the complete-case analysis, all estimates lost statistical significance; however, the confidence intervals overlapped with the ones from the primary analysis, and the direction of the associations did not change for the outcomes related to pregnancy duration (Web Tables 7 and 8). When excluding one cohort at a time, similar results were obtained. However, the odds ratios for postterm birth were unstable (Web Tables 8 and 9). In the meta-analysis stratified by timing of cell-phone use data collection, we observed that the pooled odds ratios for postterm birth in the primary analysis were driven by the DNBC and ABCD studies, which were conducted earlier and had retrospective exposure

assessment (Web Tables 8 and 10). In addition, the odds ratio for preterm birth gained statistical significance in the highly exposed group within the cohorts with prospective exposure assessment (OR = 2.03, 95% CI: 1.22, 3.39) (Web Tables 8 and 10). In the analysis with binary exposure, we observed an increased risk of giving birth at a lower gestational age (HR = 1.04, 95% CI: 1.02, 1.07) and increased odds of preterm birth (OR = 1.16, 95% CI: 1.05, 1.29) for the mothers who used their cell phones more often during pregnancy (Web Tables 8 and 11). The estimates for all of the other outcomes did not differ from unity (Web Table 11). No association of maternal cell-phone use during pregnancy with fetal growth or birth weight was detected in any of the sensitivity analyses (Web Tables 7, 9, 12, and 13). In particular, there was no association between maternal cell-phone use during pregnancy and low birth weight when the analysis was restricted to nonpreterm neonates (Web Table 13).

DISCUSSION

In this study, we examined the association of prenatal maternal cell-phone use with pregnancy duration and fetal growth outcomes in 4 general population birth cohorts. After adjusting for potential confounders, we found no association with fetal growth, but we observed an association with pregnancy duration. Women who reported more frequent calling had higher risk of giving birth at a lower gestational age compared with those reporting less frequent calling. This association was mainly driven by the preterm births; no association with postterm births was observed within the more recent cohorts (INMA and MOCEH), where postterm births were more rare. This association with pregnancy duration was reasonably stable across the cohorts, although in the Dutch cohort (ABCD) risk estimates were in the opposite direction. Notably, the association was more pronounced in the more recent cohorts (INMA and MOCEH), in which cell-phone use had been prospectively assessed during pregnancy, even though their RF-EMF exposure was expected to be lower than that in the older cohorts because of the increasing use of 3G devices.

To date, there have been few studies which examined the association of prenatal maternal cell-phone use with birth outcomes. Although our results for preterm birth are in line with those of a previous study from southern Turkey[16], an analysis of more than 100,000 births from Norway has not found such an association[18]. However, unlike our study, those studies did not control for marital status, maternal educational level, or socioeconomic position. Maternal sociodemographic characteristics correlate with cell-phone use and birth outcomes, and the direction of the association with cell-phone use has been shown to differ between populations[42-45]. Thus, residual confounding may contribute to the discrepancy between the results for preterm birth; however, it is not possible to determine the direction in which the results may have been affected.

In our study, we observed an association of maternal cell-phone use during pregnancy with pregnancy duration, but not with fetal growth. Since fetal exposure is very low during cell-phone calls[9-12], for the interpretation of these results we considered the potential effect of RF-EMF on maternal head and neck structures, as well as indirect pathways related to the use of cell phones rather than the radiation per se. Animal studies have suggested that RF-EMF exposure may result in minor thyroid gland dysfunction[14,15]. Additionally, higher preconception thyroid-stimulating hormone levels and subclinical hypothyroidism during pregnancy have been associated with higher risks of miscarriage and preterm birth[46-49]. Thus, the increased risk for giving birth preterm among heavier users of cell phones that we observed could be mediated by mild thyroid dysfunction. However, the association of RF-EMF exposure from cell-phone use with thyroid function is not established, and large-scale epidemiologic studies on the topic are lacking. Increased oxidative stress has also been considered[50]. However, it is not clear whether the elevation of radical oxygen species resulting from local RF-EMF exposure is of such an extent in humans that it could trigger systematic responses affecting birth outcomes. Causal pathways involving local radiation of parts of the human body other than the maternal head and neck were not considered, since this exposure would not be reflected in the number of cell-phone calls per day.

With regard to indirect pathways, stress may contribute to our results[51]. Psychosocial stress - acute and chronic - has been associated with higher risk for preterm birth[52-54]. Socioeconomic differences and behavioral risk factors (e.g. smoking, alcohol) contribute to this association, along with a direct biological effect[55-57]. Maternal cortisol, levels of which increase under stress, stimulates the secretion of placental corticotrophin-releasing hormone during gestation, which participates in the cascade of events initiating labor[57,58]. The elevated levels of placental corticotrophin-releasing hormone among women under stress, and to a lesser extent other stress-related mechanisms, contribute to a higher risk for preterm initiation of spontaneous labor[57,58]. Although our results were adjusted for socioeconomic position, smoking, and alcohol consumption, the direct effect of stress on pregnancy duration was not controlled for. Personal dependency and demands from work and social networks are potential sources of psychosocial stress that were not captured in the covariates used in our analyses and may correlate with cell-phone use[51].

Our study had some important strengths. The large sample size allowed us to detect potential weak associations of maternal cell-phone use with birth outcomes. All of the examined outcomes are interrelated and reflect pregnancy duration, fetal growth, and birth weight. To reduce the probability of type I error, we proposed potential pathways for statistically significant associations only if they were robust across correlated outcomes and across different cohorts. We were also able to assess whether these associations persisted or became attenuated after the introduction of 3G devices, since

our study window spanned the period during which 3G technology was introduced. Additionally, the availability of detailed information on maternal characteristics gave us the opportunity to adjust our results for confounders, which were not controlled for in previous studies.

Our study also had several limitations. The exposure variable was based only on the number of cell-phone calls per day; duration of calling was not taken into account, as that information was available only in MOCEH. Furthermore, the number of cell-phone calls per day during pregnancy was self-reported in all cohorts and was validated only in MOCEH[59]. Thus, misclassification of exposure should have attenuated the observed association, under our assumption that misclassification was predominantly nondifferential[60-62]. We expect that misclassification was much larger in the older cohorts (DNBC and ABCD), as the number of cell-phone calls per day was reported 7 years postnatal. Therefore, the estimates in the DNBC and ABCD cohorts should be more biased towards the null in comparison with the INMA and MOCEH cohorts. In addition, the etiology of the preterm births in our study population was not recorded. As a result, we could not determine whether the observed association with preterm birth was driven by spontaneous labor or labor that was induced because of pregnancy complications. Finally, maternal thyroid function during pregnancy was assessed only in a small subset of participants from the cohorts that were included in this analysis, and information about perceived stress levels during pregnancy was not consistently collected across the cohorts. Consequently, we could not explore or quantify the contributions of the proposed underlying mechanisms to the observed increase in the risk of preterm birth.

In conclusion, in our study, more frequent maternal cell-phone use during pregnancy was associated with shorter pregnancy duration, resulting in increased risk of preterm birth. No association with fetal growth or birth weight was observed. These results suggest that strong effects of cell-phone use on pregnancy duration and fetal growth are unlikely. The findings should be interpreted with caution, since they may reflect an effect of stress during pregnancy or other residual confounding rather than a direct effect of RF-EMF exposure.

Funds and financial support

GERoNiMO project: This work is supported by the European Union (grant 603794).

DNBC, Denmark: This work was supported by the Danish Epidemiology Science Centre; The Lundbeck Foundation (grant 195/04); Egmont Foundation; March of Dimes Birth Defect Foundation; Agustinus Foundation; and the Medical Research Council (grant SSVF 0646).

ABCD, The Netherlands: This work was supported by the Netherlands Organization for Health Research and Development (grant 2100.0076) and within the programme Electromagnetic Fields and Health Research (grants 85600004 and 85800001).

INMA, Spain: This work was supported by the European Union (grants FP7-ENV-2011, 282957, HEALTH.2010.2.4.5-1); Instituto de Salud Carlos III (grants G03/176, CB06/02/0041, FIS-FEDER 03/1615, 04/1509, 04/1112, 04/1931, 05/1079, 05/1052, 06/1213, 07/0314, 09/02647, 11/01007, 11/02591, CP11/00178, FIS-PI06/0867, FIS-PS09/

00090, FIS-PI041436, FIS-PI081151, FIS-PI042018, FIS-PI09/02311, FISPI13/1944, FIS-PI13/2429, FIS-PI14/0981, FIS-PI13/141687, CP13/00054 incl. FEDER funds, and MS13/00054); Conselleria de Sanitat Generalitat Valenciana; Generalitat de Catalunya (grants CIRIT1999SGR, 00241); Obra Social Cajastur; Universidad de Oviedo; Department of Health of the Basque Government (grants 2005111093, 2009111069); and the Provincial Government of Gipuzkoa (grants DFG06/004, DFG08/001).

MOCEH, Korea: This work was supported by the National Institute of Environmental Research, Ministry of Environment and the ICT R&D program of MSIP/IITP, Korea (2017-0-00961, Study on the EMF exposure control in smart society).

SUPPLEMANTARY TABLES AND FIGURES CHAPTER 5

Table S1. Birth Outcomes and maternal characteristics across exposure groups in DNBC.

	DNBC				
	Birth Outcomes				
Exposure:	None	Low	Intermediate	High	P-value
Sex (female) n(%)	14,714 (48.7)	5,274 (48.6)	3,047(49.4)	1,197 (48.8)	0.783
Birth Weight in grams m(sd)	3,615.74 (3.15)	3,595.36 (6.12)	3,565.45 (7.64)	3,565.18 (11.49)	<0.001
Low birth weight n(%)	676 (2.2)	300 (2.8)	183 (3.0)	89 (3.6)	<0.001
High birth weight n(%)	7,214 (23.9)	2,480 (22.8)	1,366 (22.1)	538 (22.0)	0.002
Gestational age at birth in weeks m(sd)	40.04 (0.010)	39.98 (0.019)	39.93 (0.023)	39.92 (0.035)	<0.001
Preterm births n(%)	1,136 (3.8)	454 (4.2)	310 (5.0)	125 (5.1)	<0.001
Postterm births n(%)	1,792 (5.9)	698 (6.4)	342 (5.5)	150 (6.1)	0.107
Birth Weight Ratio m(sd)	1.01 (0.001)	1.01 (0.001)	1.00 (0.002)	1.00 (0.003)	<0.001
Small for gestational age n(%)	1,761 (5.8)	702 (6.5)	421 (6.8)	147 (6.0)	0.008
Large for gestational age n(%)	4,747 (15.7)	1,687 (15.5)	893 (14.5)	337 (13.7)	0.008

		Maternal characteristics					Imputed values n(%)
	Exposure:	None	Low	Intermediate	High	P-value	
Age at child's birth in years m(sd)		31 (0.024)	30.27 (0.047)	29.66 (0.059)	30.2 (0.088)	<0.001	0 (0.0)
Pre-pregnancy BMI m(sd)		23.18 (0.024)	23.72 (0.046)	23.72 (0.058)	23.41 (0.086)	<0.001	2,849 (5.7)
Maternal height in cm m(sd)		168.82 (0.036)	168.91 (0.070)	168.95 (0.085)	169.26 (0.130)	0.004	2,142 (4.3)
Parity status before the index pregnancy n(%)	0	12,952 (42.9)	5,774 (53.2)	3,469 (56.2)	1,333 (54.4)	<0.001	2,131 (4.3)
	1	11,936 (39.5)	3,687 (33.9)	1,927 (31.2)	800 (32.6)		
	2	5,297 (17.5)	1,400 (12.9)	776 (12.6)	319 (13)		
Smoking during pregnancy n(%)		6,972 (23.1)	2,657 (24.5)	2,021 (32.7)	879 (35.9)	<0.001	3,689 (7.4)
Second hand smoking during pregnancy[a] n(%)		14,469 (47.9)	5,084 (46.8)	3,103 (50.3)	1,253 (51.1)	<0.001	18,870 (38)
Alcohol consumption during pregnancy n(%)		11,572 (38.3)	3,833 (35.3)	2,286 (37.0)	1,035 (42.2)	<0.001	5,061 (10.2)

Educational level, based on highest completed level n(%)	Low (primary school)	4 (0.0)	0 (0.0)	1 (0.0)	2 (0.0)	<0.001	10,967 (22.1)
	Medium (secondary school)	8,845 (29.3)	3,206 (29.5)	2,361 (38.3)	939 (38.3)		
	High (university degree)	21,336 (70.7)	7,653 (70.5)	3,810 (61.7)	1,511 (61.6)		
Socioeconomic status, based on occupational title n(%)	Low	2,089 (6.9)	795 (7.3)	633 (10.2)	294 (12.0)	<0.001	3,939 (7.9)
	Medium	10,592 (35.1)	4,005 (36.9)	2,370 (38.4)	777 (31.7)		
	High	17,504 (58.0)	6,059 (55.8)	3,170 (51.4)	1,381 (56.3)		
Marital status (unmarried) n(%)		356 (1.2)	192 (1.8)	196 (3.2)	117 (4.8)	<0.001	2,111 (4.3)
Country of birth (non-European) n(%)		487 (1.6)	157 (1.4)	129 (2.1)	48 (2.0)	0.008	115 (0.2)

[a] ≥ 1 hour per day spent in rooms that smell like tobacco smoke during pregnancy
All frequency values are averaged across the imputed datasets and rounded to nearest integer value
Abbreviations: m, mean; n, sample size; sd, standard deviation

Table S2. Birth Outcomes and maternal characteristics across exposure groups in ABCD.

		ABCD					
		Birth Outcomes					
	Exposure:	None	Low	Intermediate	High	P-value	
Sex (female) n(%)		96 (53.3)	532 (47.3)	337 (47.9)	300 (50.9)	0,29	
Birth Weight in grams m(sd)		3,478.97 (38.85)	3,504.23 (41.84)	3,509.8 (43.54)	3,500.68 (44.39)	0,91	
Low birth weight n(%)		6 (3.3)	35 (3.1)	21 (3.0)	15 (2.5)	0,91	
High birth weight n(%)		28 (15.6)	202 (18.0)	106 (15.1)	88 (14.9)	0,27	
Gestational age at birth in weeks m(sd)		39.81 (0.120)	39.91 (0.129)	39.93 (0.135)	40.01 (0.137)	0,46	
Preterm births n(%)		6 (3.3)	54 (4.8)	33 (4.7)	24 (4.1)	0,77	
Postterm births n(%)		7 (3.9)	50 (4.4)	24 (3.4)	27 (4.6)	0,68	
Birth Weight Ratio m(sd)		1.00 (0.009)	1.01 (0.010)	1.02 (0.010)	1.01 (0.011)	0,44	
Small for gestational age n(%)		15 (8.3)	95 (8.4)	50 (7.1)	46 (7.8)	0,78	
Large for gestational age n(%)		18 (10.0)	125 (11.1)	75 (10.7)	58 (9.8)	0,87	
		Maternal characteristics					Imputed values n(%)
	Exposure:	None	Low	Intermediate	High	P-value	
Age at child's birth in years m(sd)		33.78 (0.318)	33.36 (0.345)	32.54 (0.357)	33.17 (0.364)	<0.001	147 (5.7)
Pre-pregnancy BMI m(sd)		23.58 (0.269)	22.86 (0.290)	22.53 (0.301)	22.25 (0.307)	<0.001	2 (0.1)
Maternal height in cm m(sd)		169.29 (0.498)	169.71 (0.536)	170.99 (0.558)	170.78 (0.569)	<0.001	0 (0.0)
Parity status before the index pregnancy n(%)	0,00	68 (37.8)	605 (53.8)	476 (67.7)	372 (63.2)	<0.001	0(0.0)
	1,00	75 (41.7)	415 (36.9)	188 (26.7)	167 (28.4)		
	2,00	37 (20.6)	105 (9.3)	39 (5.5)	50 (8.5)		
Smoking during pregnancy n(%)		4 (2.2)	70 (6.2)	60 (8.5)	64 (10.9)	<0.001	0 (0.0)
Second hand smoking during pregnancy n(%)		63 (35.0)	538 (47.8)	434 (61.7)	372 (63.2)	<0.001	3 (0.1)
Alcohol consumption during pregnancy n(%)		25 (13.9)	262 (23.3)	231 (32.9)	224 (38.0)	<0.001	0 (0.0)
Educational level, based on years of education after primary school n(%)	Low (≤5)	41 (22.8)	125 (11.1)	40 (5.7)	39 (6.6)	<0.001	9 (0.3)
	Medium (6-9)	36 (19.8)	258 (22.9)	168 (23.8)	109 (18.5)		
	High (≥10)	103 (57.4)	742 (66.0)	495 (70.5)	441 (74.9)		
Socioeconomic status, based on ISCO88 categories n(%)	Low (I/II)	11 (6.2)	28 (2.5)	3 (0.4)	8 (1.3)	<0.001	621 (23.9)
	Medium (III)	52 (28.9)	276 (24.5)	126 (18.0)	81 (13.8)		
	High (IV/V)	117 (64.9)	822 (73.0)	574 (81.7)	500 (84.9)		
Marital status (not cohabitating with the father) n(%)		16 (8.9)	100 (8.9)	50 (7.1)	55 (9.3)	0,48	1 (0.0)
Country of birth (non-European) n(%)		48 (26.7)	149 (13.2)	44 (6.3)	45 (7.6)	<0.001	0 (0.0)

All frequency values are averaged across the imputed datasets and rounded to nearest integer value
Abbreviations: m, mean; n, sample size; sd, standard deviation

Table S3. Birth Outcomes and maternal characteristics across exposure groups in INMA.

		INMA					
		Birth Outcomes					
	Exposure:	None	Low	Intermediate	High	P-value	
Sex (female) n(%)		25 (47.2)	333 (47.4)	370 (49.1)	211 (49.6)	0,864	
Birth Weight in grams m(sd)		3333.11 (63.41)	3272.27 (65.76)	3244.13 (65.60)	3271.09 (67.25)	0,408	
Low birth weight n(%)		2 (3.8)	31 (4.4)	34 (4.5)	22 (5.2)	0,925	
High birth weight n(%)		1 (1.9)	43 (6.1)	37 (4.9)	27 (6.4)	0,412	
Gestational age at birth in weeks m(sd)		39.93 (0.219)	39.83 (0.228)	39.88 (0.227)	39.75 (0.233)	0,56	
Preterm births n(%)		2 (3.8)	22 (3.1)	22 (2.9)	19 (4.5)	0,532	
Postterm births n(%)		0 (0.0)	21 (3.0)	30 (4.0)	21 (4.9)	0,171	
Birth Weight Ratio m(sd)		1.03 (0.017)	1.01 (0.018)	1.00 (0.018)	1.02 (0.018)	0,176	
Small for gestational age n(%)		2 (3.8)	58 (8.3)	85 (11.3)	37 (8.7)	0,091	
Large for gestational age n(%)		6 (11.3)	75 (10.7)	73 (9.7)	51 (12.0)	0,668	
		Maternal characteristics					Imputed values n(%)
	Exposure:	None	Low	Intermediate	High	P-value	
Age at child's birth in years m(sd)		32.33 (0.587)	32.2 (0.609)	31.57 (0.607)	31.24 (0.622)	0,001	79 (4.1)
Pre-pregnancy BMI m(sd)		24.12 (0.592)	23.57 (0.614)	23.44 (0.612)	23.55 (0.627)	0,718	5 (0.3)
Maternal height in cm m(sd)		162.57 (0.846)	162.5 (0.877)	162.74 (0.875)	163.33 (0.897)	0,175	1 (0.1)
Parity status before the index pregnancy n(%)	0	20 (37.7)	384 (54.6)	427 (56.7)	241 (56.7)	0,009	2 (0.1)
	1	31 (58.5)	254 (36.1)	284 (37.7)	157 (36.9)		
	2	2 (3.8)	65 (9.2)	42 (5.6)	27 (6.4)		
Smoking during pregnancy n(%)		14 (26.4)	198 (28.2)	262 (34.8)	155 (36.4)	0,009	4 (0.2)
Second hand smoking during pregnancy[a] n(%)		31 (58.5)	417 (59.3)	516 (68.5)	308 (72.4)	<0.001	16 (0.8)
Alcohol consumption during pregnancy n(%)		3 (5.8)	66 (9.3)	79 (10.5)	36 (8.4)	0,518	21 (1.1)
Educational level, based on highest completed level n(%)	Low (primary school)	17 (32.1)	188 (26.7)	207 (27.5)	87 (20.5)	0,028	3 (0.2)
	Medium (secondary school)	21 (39.6)	294 (41.9)	305 (40.5)	168 (39.4)		
	High (university degree)	15 (28.3)	221 (31.4)	241 (32.0)	170 (40.0)		
Socioeconomic status, based on occupational title n(%)	Low	30 (56.6)	375 (53.3)	414 (55.0)	181 (42.6)	0,002	0 (0.0)
	Medium	10 (18.9)	190 (27.0)	189 (25.1)	135 (31.8)		
	High	13 (24.5)	138 (19.6)	150 (19.9)	109 (25.6)		
Marital status (not cohabitating with the father) n(%)		1 (1.9)	10 (1.4)	14 (1.9)	5 (1.2)	0,808	0 (0.0)
Country of birth (non-European) n(%)		9 (17.0)	47 (6.7)	42 (5.6)	29 (6.8)	0,014	0 (0.0)

[a] Smoker at home during pregnancy
All frequency values are averaged across the imputed datasets and rounded to nearest integer value
Abbreviations: m, mean; n, sample size; sd, standard deviation

Table S4. Birth Outcomes and maternal characteristics across exposure groups in MOCEH.

		MOCEH					
		Birth Outcomes					
	Exposure:	**None**	**Low**	**Intermediate**	**High**	**P-value**	
Sex (female) n(%)		6 (40.0)	119 (49.2)	293 (45.6)	206 (50.4)	0,42	
Birth Weight in grams m(sd)		3385.33 (113.82)	3266.22 (117.30)	3278.51 (115.15)	3227.14 (115.89)	0,199	
Low birth weight n(%)		0 (0.0)	7 (2.9)	13 (2.0)	14 (3.4)	0,491	
High birth weight n(%)		1 (6.7)	14 (5.8)	34 (5.3)	9 (2.2)	0,066	
Gestational age at birth in weeks m(sd)		39.5 (0.413)	39.27 (0.426)	39.23 (0.418)	39.06 (0.421)	0,231	
Preterm births n(%)		1 (6.7)	9 (3.7)	28 (4.4)	26 (6.4)	0,378	
Postterm births n(%)		0 (0.0)	1 (0.4)	4 (0.6)	3 (0.7)	0,95	
Birth Weight Ratio m(sd)		1.03 (0.030)	1.00 (0.031)	1.01 (0.031)	1.00 (0.031)	0,843	
Small for gestational age n(%)		1 (6.7)	22 (9.1)	52 (8.1)	42 (10.3)	0,673	
Large for gestational age n(%)		2 (13.3)	29 (12.0)	71 (11.1)	44 (10.8)	0,958	
		Maternal characteristics					Imputed values n(%)
	Exposure:	**None**	**Low**	**Intermediate**	**High**	**P-value**	
Age at child's birth in years m(sd)		30.9 (0.954)	31.22 (0.983)	30.67 (0.965)	30.78 (0.971)	0,264	0 (0.0)
Pre-pregnancy BMI m(sd)		22.55 (0.755)	20.92 (0.778)	21.21 (0.764)	21.31 (0.769)	0,115	27 (2.1)
Maternal height in cm m(sd)		161.27 (1.224)	161.46 (1.262)	161.08 (1.239)	161.29 (1.247)	0,736	14 (1.1)
Parity status before the index pregnancy n(%)	0	7 (45.0)	96 (39.6)	332 (51.7)	219 (53.5)	0,005	172 (13.1)
	1	5 (34.7)	125 (51.5)	251 (39.1)	150 (36.6)		
	2	3 (20.3)	22 (8.9)	59 (9.1)	40 (9.9)		
Smoking during pregnancy n(%)		0 (0.0)	2 (0.9)	1 (0.2)	6 (1.5)	0,129	27 (2.1)
Second hand smoking during pregnancy[a] n(%)		8 (53.3)	143 (59.1)	404 (62.9)	250 (61.1)	0,662	0 (0.0)
Alcohol consumption during pregnancy n(%)		2 (11.3)	13 (5.4)	37 (5.8)	28 (6.9)	0,775	114 (8.7)
Educational level, based on highest completed level n(%)	Low (primary school)	1 (6.7)	1 (0.4)	0 (0.0)	3 (0.7)	0,024	16 (1.2)
	Medium (secondary school)	6 (40.0)	62 (25.5)	177 (27.6)	101 (24.7)		
	High (university degree)	8 (53.3)	179 (74.1)	465 (72.4)	305 (74.5)		
Socioeconomic status n(%) Income level KRW*10,000	50-100	0 (0.0)	2 (0.8)	5 (0.8)	2 (0.5)	0,002	32 (2.4)
	100-150	4 (26.7)	12 (5.0)	56 (8.7)	33 (8.1)		
	150-200	3 (20.0)	41 (16.9)	136 (21.2)	60 (14.7)		
	200-300	6 (40.0)	96 (39.7)	213 (33.2)	126 (30.8)		
	300-400	0 (0.0)	48 (19.8)	130 (20.2)	82 (20.0)		
	400-600	2 (13.3)	34 (14.0)	86 (13.4)	78 (19.1)		
	≥600	0 (0.0)	10 (4.1)	16 (2.5)	28 (6.8)		
Marital status (not cohabitating with the father) n(%)		0 (0.0)	4 (1.7)	12 (1.9)	9 (2.2)	0,943	23 (1.8)

[a] Smoking at home during pregnancy
All frequency values are averaged across the imputed datasets and rounded to nearest integer value. Abbreviations: KRW, South Korean Won; m, mean; n, sample size; sd, standard deviation

Table S5. Meta-analysis of unadjusted estimates for the effect of maternal cell phone use during pregnancy on birth outcomes.

	Exposure	None	Low	Intermediate	High
Birth weight	*Outcome: Birth weight within non-preterm neonates (Mean Difference in grams)*				
	DNBC	**12.30 (1.26, 23.34)**	ref	**-24.09 (-39.86, -8.31)**	-17.82 (-39.98, 4.34)
	ABCD	-39.86 (-116.72, 36.99)	ref	6.90 (-39.41, 53.22)	-8.75 (-57.64, 40.14)
	INMA	47.89 (-74.48, 170.26)	ref	-31.73 (-76.62,13.16)	17.61 (-35.24, 70.46)
	MOCEH	62.90 (-143.25, 269.04)	ref	22.78 (-34.86, 80.42)	-10.52 (-72.76, 51.72)
	Pooled MD	**11.68 (0.81, 22.54)**	ref	-16.20 (-35.13, 2.73)	-11.84 (-29.88, 6.20)
	Outcome: Low birth weight (Odds Ratio)				
	DNBC	**0.81 (0.70, 0.93)**	ref	1.08 (0.89, 1.30)	**1.33 (1.04, 1.69)**
	ABCD	1.07 (0.45, 2.59)	ref	0.96 (0.55, 1.66)	0.81 (0.44, 1.50)
	INMA	0.85 (0.20, 3.65)	ref	1.03 (0.62, 1.69)	1.18 (0.68, 2.07)
	MOCEH	0.00 (0.00, Inf)	ref	0.69 (0.27, 1.76)	1.19 (0.47, 2.99)
	Pooled OR	**0.81 (0.71, 0.93)** [a]	ref	1.04 (0.89, 1.23)	**1.23 (1.01, 1.51)**
	Outcome: High birth weight (Odds Ratio)				
	DNBC	**1.06 (1.01, 1.12)**	ref	0.96 (0.89, 1.04)	0.95 (0.86, 1.06)
	ABCD	0.84 (0.55, 1.30)	ref	0.81 (0.63, 1.05)	0.80 (0.61, 1.05)
	INMA	0.30 (0.04, 2.19)	ref	0.79 (0.50, 1.25)	1.04 (0.63, 1.71)
	MOCEH	1.16 (0.14, 9.49)	ref	0.91 (0.48, 1.73)	**0.37 (0.16, 0.86)**
	Pooled OR	1.02 (0.87, 1.20)	ref	0.92 (0.82, 1.03)	0.89 (0.77, 1.04)
Pregnancy duration	*Outcome : Gestational age at birth (Hazard ratio)*				
	DNBC	1.00 (0.98, 1.02)	ref	**1.04 (1.01, 1.08)**	1.03 (0.98, 1.07)
	ABCD	1.09 (0.93, 1.27)	ref	1.00 (0.91, 1.10)	0.97 (0.87, 1.07)
	INMA	0.97 (0.73, 1.28)	ref	0.93 (0.84, 1.03)	0.99 (0.88, 1.12)
	MOCEH	0.87 (0.51, 1.46)	ref	1.02 (0.88, 1.18)	1.02 (0.87, 1.20)
	Pooled HR	1.00 (0.98, 1.02)	ref	1.01 (0.96, 1.06)	1.01 (0.98, 1.05)
	Outcome: Preterm birth(Odds Ratio)				
	DNBC	0.90 (0.80, 1.00)	ref	**1.21 (1.05, 1.41)**	**1.23 (1.01, 1.51)**
	ABCD	0.68 (0.29, 1.61)	ref	0.98 (0.63, 1.52)	0.84 (0.52, 1.38)
	INMA	1.21 (0.28, 5.31)	ref	0.93 (0.51, 1.70)	1.45 (0.77, 2.71)
	MOCEH	1.85 (0.22, 15.64)	ref	1.18 (0.55, 2.54)	1.76 (0.81, 3.82)
	Pooled OR	**0.90 (0.80, 1.00)**	ref	**1.17 (1.02, 1.34)**	**1.21 (1.02, 1.44)**
	Outcome: Postterm birth(Odds Ratio)				
	DNBC	0.92 (0.84, 1.01)	ref	**0.85 (0.75, 0.98)**	0.95 (0.79, 1.14)
	ABCD	0.87 (0.39, 1.95)	ref	0.76 (0.46, 1.25)	1.03 (0.64, 1.67)
	INMA	0.00 (0.00, Inf)	ref	1.35 (0.76, 2.38)	1.69 (0.91, 3.13)
	MOCEH	0.00 (0.0, Inf)	ref	1.51 (0.17, 13.59)	1.78 (0.18, 17.22)
	Pooled OR	0.92 (0.84, 1.00) [b]	ref	**0.87 (0.77, 0.98)**	1.06 (0.82, 1.37)

	Exposure	None	Low	Intermediate	High
Fetal growth	*Outcome: Birth weight ratio (Mean difference)*				
	DNBC	**0.00 (0.00,0.01)**	ref	**-0.01 (-0.01,-0.00)**	-0.00 (-0.01,0.00)
	ABCD	-0.01 (-0.03,0.01)	ref	0.01 (-0.00,0.02)	0.00 (-0.01,0.01)
	INMA	0.02 (-0.02, 0.05)	ref	-0.01 (-0.02,0.00)	0.01 (-0.01, 0.02)
	MOCEH	0.02 (-0.04, 0.08)	ref	0.00 (-0.01, 0.02)	-0.00 (-0.02, 0.02)
	Pooled MD	**0.00 (0.00,0.01)**	ref	-0.00 (-0.01,0.01)	-0.00 (-0.01, 0.00)
	Outcome: Small for gestational age (Odds Ratio)				
	DNBC	**0.90 (0.82, 0.98)**	ref	1.06 (0.93, 1.20)	0.92 (0.77, 1.11)
	ABCD	0.99 (0.56, 1.74)	ref	0.83 (0.58, 1.19)	0.92 (0.64, 1.33)
	INMA	0.44 (0.10, 1.84)	ref	1.42 (1.00, 2.01)	1.06 (0.69, 1.63)
	MOCEH	0.71 (0.09, 5.69)	ref	0.88 (0.52, 1.49)	1.14 (0.67, 1.97)
	Pooled OR	**0.90 (0.82, 0.98)**	ref	1.05 (0.87, 1.26)	0.95 (0.82, 1.10)
	Outcome: Large for gestational age (Odds Ratio)				
	DNBC	1.01 (0.96, 1.08)	ref	0.92 (0.84, 1.00)	**0.87 (0.76, 0.98)**
	ABCD	0.89 (0.53, 1.50)	ref	0.96 (0.71, 1.29)	0.87 (0.63, 1.21)
	INMA	1.07 (0.44, 2.58)	ref	0.90 (0.64, 1.26)	1.14 (0.78, 1.67)
	MOCEH	1.13 (0.24, 5.26)	ref	0.91 (0.58, 1.45)	0.89 (0.54, 1.46)
	Pooled OR	1.01 (0.95, 1.08)	ref	**0.92 (0.85, 1.00)**	**0.89 (0.80, 0.99)**

[a] excluding MOCEH
[b] excluding MOCEH and INMA
Abbreviations: HR, hazard ratio; Inf, infinity; MD, mean difference; OR, odds ratio; ref, reference level
Statistically significant results marked with bold

Table S6. Meta-analysis of adjusted estimates for the effect of maternal cell phone use during pregnancy on birth outcomes.

	Exposure	None	Low	Intermediate	High
Birth weight	*Outcome: Birth weight within non-preterm neonates (Mean Difference in grams)*				
	DNBC	4.83(-5.70,15.36)	ref	-8.68(-23.71,6.35)	-1.93(-23.08,19.31)
	ABCD	-65.88 (-139.49, 7.73)	ref	2.68 (-42.17, 47.53)	-3.26 (-50.52,44.01)
	INMA	0.19(-117.73,118.11)	ref	-28.19(-71.69,15.31)	9.41(-42.24,61.05)
	MOCEH	4.21(-198.34,206.77)	ref	15.60(-40.91,72.10)	-23.48(-84.67,37.70)
	Pooled MD	-11.15(-53.24,30.94)	ref	-8.17(-21.34,5.00)	-2.56(-19.90,14.78)
	Outcome: Low birth weight (Odds Ratio)				
	DNBC	**0.86(0.75,0.99)**	ref	0.95(0.79,1.15)	1.13(0.89,1.45)
	ABCD	1.12(0.45,2.79)	ref	1.00(0.56,1.76)	0.86(0.46,1.62)
	INMA	1.08(0.24,4.80)	ref	1.02(0.61,1.71)	1.41(0.78,2.53)
	MOCEH	0.00(0.00,Inf)	ref	0.64(0.24,1.68)	1.24(0.47,3.28)
	Pooled OR	**0.87 (0.76, 1.00)** [a]	ref	0.95(0.81,1.13)	1.13(0.92,1.40)
	Outcome: High birth weight (Odds Ratio)				
	DNBC	1.03(0.98,1.09)	ref	1.01(0.94,1.10)	1.01(0.90,1.12)
	ABCD	0.74(0.47,1.16)	ref	0.83(0.63,1.09)	0.83(0.62,1.11)
	INMA	0.28(0.04,2.11)	ref	0.80(0.50,1.28)	1.05(0.63,1.78)
	MOCEH	1.14(0.13,10.41)	ref	1.03(0.51,2.10)	**0.37(0.15,0.92)**
	Pooled OR	0.92(0.68,1.24)	ref	0.96(0.83,1.10)	0.93(0.78,1.11)

	Exposure	None	Low	Intermediate	High
Pregnancy duration	*Outcome : Gestational age at birth (Hazard ratio)*				
	DNBC	0.99 (0.97,1.01)	ref	**1.04 (1.01, 1.08)**	1.02 (0.98,1.07)
	ABCD	1.06 (0.90,1.25)	ref	1.07 (0.97, 1.18)	0.99 (0.89,1.10)
	INMA	0.94 (0.70,1.26)	ref	0.95 (0.85,1.06)	1.01 (0.89,1.16)
	MOCEH	0.85 (0.48,1.51)	ref	1.08 (0.92,1.27)	1.10 (0.93,1.31)
	Pooled HR	0.99 (0.97,1.01)	ref	**1.04 (1.01,1.07)**	1.02 (0.98,1.06)
	Outcome: Preterm birth(Odds Ratio)				
	DNBC	0.96(0.86,1.07)	ref	1.13(0.98,1.32)	1.14(0.93,1.40)
	ABCD	0.76(0.32,1.82)	ref	0.91(0.58,1.44)	0.83(0.50,1.38)
	INMA	1.48(0.32,6.89)	ref	1.16(0.62,2.17)	**2.16(1.11,4.21)**
	MOCEH	1.82(0.20,16.87)	ref	1.29(0.59,2.82)	1.85(0.83,4.12)
	Pooled OR	0.96(0.86,1.07)	ref	1.12(0.97,1.28)	1.28(0.87,1.88)
	Outcome: Postterm birth(Odds Ratio)				
	DNBC	0.98(0.89,1.07)	ref	**0.85(0.74,0.97)**	0.95(0.79,1.14)
	ABCD	1.02(0.44,2.35)	ref	0.65(0.39,1.09)	0.92(0.56,1.51)
	INMA	0.00(0.00,Inf)	ref	1.32(0.74,2.36)	1.61(0.85,3.07)
	MOCEH	0.00(0.00,Inf)	ref	1.40(0.13,14.91)	1.34(0.11,16.50)
	Pooled OR	0.98 (0.89, 1.07) [b]	ref	**0.85(0.75,0.97)**	0.98(0.83,1.16)
Fetal growth	*Outcome: Birth weight ratio (Mean difference)*				
	DNBC	0.00(-0.00,0.00)	ref	-0.00(-0.01,0.00)	0.00(-0.01,0.01)
	ABCD	-0.01(-0.03,0.01)	ref	0.00(-0.01,0.01)	-0.00(-0.01,0.01)
	INMA	0.00(-0.03,0.04)	ref	-0.01(-0.02,0.01)	0.00(-0.01,0.02)
	MOCEH	0.01(-0.05,0.07)	ref	0.00(-0.01,0.02)	-0.00(-0.02,0.02)
	Pooled MD	0.00(-0.00,0.00)	ref	-0.00(-0.00,0.00)	0.00(-0.00,0.00)
	Outcome: Small for gestational age (Odds Ratio)				
	DNBC	0.94(0.86,1.03)	ref	0.97(0.86,1.11)	0.83(0.69,1.00)
	ABCD	0.85(0.47,1.54)	ref	1.01(0.70,1.48)	1.01(0.69,1.50)
	INMA	0.56(0.13,2.41)	ref	1.41(0.98,2.03)	1.10(0.70,1.73)
	MOCEH	0.82(0.10,6.90)	ref	0.86(0.50,1.48)	1.19(0.68,2.10)
	Pooled OR	0.94(0.86,1.03)	ref	1.03(0.88,1.21)	0.94(0.78,1.13)
	Outcome: Large for gestational age (Odds Ratio)				
	DNBC	0.98(0.92,1.05)	ref	0.97(0.89,1.07)	0.92(0.80,1.05)
	ABCD	0.79(0.46,1.37)	ref	0.93(0.67,1.27)	0.88(0.62,1.24)
	INMA	0.93(0.38,2.33)	ref	0.93(0.65,1.33)	1.13(0.76,1.68)
	MOCEH	0.94(0.19,4.68)	ref	0.93(0.57,1.53)	0.89(0.52,1.53)
	Pooled OR	0.98(0.92,1.04)	ref	0.97(0.89,1.05)	0.93(0.83,1.04)

Abbreviations: HR, hazard ratio; Inf, infinity; MD, mean difference; OR, odds ratio; ref, reference level
[a] excluding MOCEH, [b] excluding MOCEH and INMA
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Statistically significant results marked with bold

Table S7. Meta-analysis of adjusted estimates from complete case analyses for the effect of cell mobile phone use during pregnancy on birth outcomes.

	Exposure	None	Low	Intermediate	High
Birth weight	*Outcome: Birth weight within non-preterm neonates (Mean Difference in grams)*				
	DNBC	9.27 (-5.84,24.38)	ref	4.46 (-16.36,25.29)	7.92 (-22.58,38.42)
	ABCD	-69.17 (-165.72,27.39)	ref	-22.48 (-75.02,30.07)	-9.79 (-64.79,45.21)
	INMA	-20.68 (-142.09,100.74)	ref	-32.76 (-77.63, 12.11)	-0.12 (-53.15,52.90)
	MOCEH	100.24 (-148.85,349.33)	ref	9.86 (-55.62,75.34)	-23.66 (-94.51,47.18)
	Pooled MD	-2.12 (-41.32,37.07)	ref	-5.76 (-26.59,15.07)	0.27 (-22.32,22.85)
	Outcome: Low birth weight (Odds Ratio)				
	DNBC	0.90 (0.71, 1.13)	ref	0.96 (0.71,1.30)	0.90 (0.57,1.41)
	ABCD	1.32 (0.44, 4.02)	ref	1.13 (0.58,2.19)	0.98 (0.46,2.05)
	INMA	1.10 (0.25, 4.95)	ref	0.84 (0.49,1.44)	1.25 (0.68,2.30)
	MOCEH	0.00 (0.00,Inf)	ref	0.55 (0.15,2.02)	0.92 (0.25,3.41)
	Pooled OR	0.91 (0.73, 1.14)[a]	ref	0.94 (0.74,1.19)	1.00 (0.73,1.37)
	Outcome: High birth weight (Odds Ratio)				
	DNBC	1.04 (0.96,1.13)	ref	1.05 (0.94,1.17)	1.05 (0.90,1.23)
	ABCD	0.69 (0.38,1.25)	ref	0.72 (0.52,1.00)	0.76 (0.54,1.07)
	INMA	0.28 (0.04,2.10)	ref	0.67 (0.41,1.10)	1.02 (0.60,1.74)
	MOCEH	1.68 (0.17,16.85)	ref	1.37 (0.61,3.07)	0.40 (0.14,1.13)
	Pooled OR	0.92 (0.65, 1.30)	ref	0.89 (0.68,1.17)	0.91 (0.71,1.16)
Pregnancy duration	*Outcome : Gestational age at birth (Hazard ratio)*				
	DNBC	1.00 (0.14,7.07)	ref	1.03 (0.14,7.83)	1.01 (0.14,7.29)
	ABCD	0.94 (0.15,5.87)	ref	1.09 (0.13,9.35)	1.03 (0.14,7.84)
	INMA	0.97 (0.15,6.44)	ref	0.92 (0.15,5.56)	1.02 (0.14,7.59)
	MOCEH	0.84 (0.16,4.36)	ref	1.05 (0.13,8.23)	1.09 (0.13,9.28)
	Pooled HR	0.93 (0.37,2.30)	ref	1.01 (0.37,2.75)	1.04 (0.38,2.87)
	Outcome: Preterm birth(Odds Ratio)				
	DNBC	0.93 (0.78,1.11)	ref	1.13 (0.90,1.42)	0.76 (0.52,1.13)
	ABCD	0.83 (0.28,2.40)	ref	0.80 (0.47,1.36)	0.68 (0.37,1.24)
	INMA	1.68 (0.35,7.94)	ref	1.03 (0.52,2.02)	1.97 (0.96,4.02)
	MOCEH	2.12 (0.21,21.67)	ref	1.43 (0.58,3.54)	1.53 (0.59,3.96)
	Pooled OR	0.94 (0.79,1.12)	ref	1.08 (0.89,1.32)	1.03 (0.62,1.72)
	Outcome: Postterm birth(Odds Ratio)				
	DNBC	0.92 (0.80,1.05)	ref	0.84 (0.70,1.01)	0.89 (0.67,1.17)
	ABCD	1.32 (0.53,3.31)	ref	**0.52 (0.29,0.96)**	0.73 (0.41,1.32)
	INMA	0.00 (0.00,Inf)	ref	1.58 (0.85,2.92)	1.61 (0.80, 3.24)
	MOCEH	0.00 (0.00,Inf)	ref	1.60 (0.14,17.55)	1.49 (0.11,19.98)
	Pooled OR	0.93 (0.81, 1.06)[b]	ref	0.90 (0.54,1.50)	0.92 (0.73,1.17)

	Exposure	None	Low	Intermediate	High
Fetal growth	*Outcome: Birth weight ratio (Mean difference)*				
	DNBC	0.00 (-0.00,0.01)	ref	0.00 (-0.00,0.01)	0.00 (-0.01,0.01)
	ABCD	-0.02 (-0.04,0.01)	ref	-0.00 (-0.02,0.01)	-0.00 (-0.02,0.01)
	INMA	0.00 (-0.03,0.03)	ref	-0.01 (-0.02,0.01)	0.00 (-0.01,0.02)
	MOCEH	0.03 (-0.04,0.11)	ref	0.00 (-0.02, 0.02)	0.00 (-0.02,0.02)
	Pooled MD	-0.00 (-0.01,0.01)	ref	-0.00 (-0.00,0.00)	0.00 (-0.01,0.01)
	Outcome: Small for gestational age (Odds Ratio)				
	DNBC	0.93 (0.81,1.07)	ref	0.90 (0.75,1.09)	0.80 (0.60,1.07)
	ABCD	0.75 (0.31,1.86)	ref	1.29 (0.83,2.01)	1.25 (0.78,1.99)
	INMA	0.61 (0.14,2.64)	ref	1.39 (0.95,2.03)	1.14 (0.71,1.81)
	MOCEH	1.59 (0.17,14.67)	ref	1.25 (0.63,2.50)	1.24 (0.59,2.60)
	Pooled OR	0.93 (0.81,1.06)	ref	1.13 (0.87,1.47)	1.01 (0.77,1.33)
	Outcome: Large for gestational age (Odds Ratio)				
	DNBC	0.99 (0.90,1.08)	ref	1.01 (0.89,1.15)	0.87 (0.72,1.06)
	ABCD	0.74 (0.36,1.50)	ref	0.86 (0.59,1.25)	0.96 (0.65,1.41)
	INMA	1.00 (0.40,2.50)	ref	0.84 (0.58,1.22)	1.07 (0.71,1.63)
	MOCEH	1.44 (0.27,7.76)	ref	1.09 (0.61,1.94)	0.89 (0.48,1.66)
	Pooled OR	0.99 (0.90,1.08)	ref	0.98 (0.88,1.10)	0.91 (0.78,1.06)

Abbreviations: HR, hazard ratio; Inf, infinity; MD, mean difference; OR, odds ratio; ref, reference level
[a]excluding MOCEH, [b]excluding MOCEH and INMA
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Statistically significant results marked with bold

Table S8. Summary of sensitivity analyses for pregnancy duration outcomes.

Sensitivity Analsysis		Exposure			
		None	Low	Intermediate	High
Complete case analysis					
Gestational age at birth (HR)		0.93 (0.37,2.30)	ref	1.01 (0.37,2.75)	1.04 (0.38,2.87)
Preterm birth (OR)		0.94 (0.79,1.12)	ref	1.08 (0.89,1.32)	1.03 (0.62,1.72)
Postterm birth (OR)		0.93 (0.81, 1.06)*	ref	0.90 (0.54,1.50)	0.92 (0.73,1.17)
Analysis with binary exposure					
Gestational age at birth (HR)		ref		**1.04 (1.02,1.07)**	
Preterm birth (OR)		ref		**1.16 (1.05,1.29)**	
Postterm birth (OR)		ref		0.95 (0.73, 1.23)	
Excluding one cohort					
Gestational age at birth (HR)	excluding DNBC	1.02 (0.89,1.17)	ref	1.02 (0.94,1.12)	1.02 (0.94,1.09)
	excluding ABCD	0.99 (0.97,1.01)	ref	1.02 (0.97, 1.09)	1.03 (0.99, 1.07)
	excluding INMA	0.99 (0.97 1.01)	ref	**1.05 (1.02, 1.08)**	1.02 (0.98, 1.06)
	excluding MOCEH	0.99 (0.97, 1.01)	ref	**1.04 (1.00, 1.07)**	1.02 (0.98, 1.06)
Preterm birth (OR)	excluding DNBC	0.96 (0.47,1.98)	ref	1.04 (0.75,1.45)	1.43 (0.76,2.68)
	excluding ABCD	0.96 (0.86, 1.08)	ref	1.14 (0.99, 1.31)	1.49 (0.95, 2.33)
	excluding INMA	0.96 (0.86,1.07)	ref	1.11 (0.97, 1.28)	1.12(0.93, 1.35)
	excluding MOCEH	0.96 (0.86, 1.07)	ref	1.11 (0.97, 1.28)	1.20 (0.78, 1.87)
Postterm birth (OR)	excluding DNBC	1.02(0.44,2.35)	ref	0.94 (0.51,1.74)	1.17 (0.71,1.95)
	excluding ABCD	0.98(0.89,1.07)	ref	0.97 (0.67, 1.40)	1.12 (0.71, 1.77)
	excluding INMA	-	ref	**0.84 (0.73, 0.95)**	0.95 (0.80, 1.13)
	excluding MOCEH	-	ref	**0.85 (0.75, 0.97)**	0.98 (0.83, 1.16)
Timing of cell phone use data collection					
Gestational age at birth (HR)					
Retrospectively:	DNBC & ABCD	0.99 (0.97,1.01)	ref	**1.04 (1.01,1.08)**	1.02 (0.98,1.06)
Prospectively:	INMA & MOCEH	0.92 (0.71,1.19)	ref	1.00 (0.88,1.13)	1.05 (0.94,1.16)
Preterm birth (OR)					
Retrospectively:	DNBC & ABCD	0.96 (0.86, 1.07)	ref	1.11 (0.96,1.28)	1.06 (0.82,1.38)
Prospectively:	INMA & MOCEH	1.58 (0.45,5.61)	ref	1.21 (0.74, 1.97)	**2.03 (1.22,3.39)**
Postterm birth (OR)					
Retrospectively:	DNBC & ABCD	0.98 (0.89, 1.07)	ref	**0.83 (0.73,0.95)**	0.95 (0.80,1.12)
Prospectively:	INMA & MOCEH	0.00 (0.00,Inf)	ref	1.32 (0.75,2.32)	1.60 (0.86,2.97)

*excluding INMA & MOCEH

Abbreviations: HR, hazard ratio; MD, mean difference; OR, odds ratio; ref, reference level. Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Statistically significant results marked with bold

Table S9. Meta-analysis of adjusted estimates for the effect of maternal cell phone use during pregnancy on birth outcomes, excluding one cohort at a time.

	Exposure	None	Low	Intermediate	High
Birth weight	Excluded cohort	*Outcome: Birth weight within non-preterm neonates (Mean Difference in grams)*			
	DNBC	-42.88 (-102.55,16.79)	ref	-6.49 (-33.88,20.84)	-3.86 (-34.15,26.44)
	ABCD	4.79 (-5.69,15.26)	ref	-9.19 (-22.97,4.59)	-2.46 (-21.09,16.18)
	INMA	-16.58 (-72.08, 38.92)	ref	-6.15 (-19.97, 7.67)	-4.08 (-22.49, 14.32)
	MOCEH	-13.30 (-60.02, 33.41)	ref	-9.53 (-23.08, 4.01)	-0.74 (-18.82, 17.34)
	Pooled MD	-11.15(-53.24,30.94)	ref	-8.17(-21.34,5.00)	-2.56(-19.90,14.78)
	Excluded cohort	*Outcome: Low birth weight (Odds Ratio)*			
	DNBC	-	ref	0.95 (0.67,1.36)	1.14 (0.77,1.69)
	ABCD	-	ref	0.95 (0.80, 1.13)	1.17 (0.94, 1.46)
	INMA	-	ref	0.94 (0.79,1.13)	1.10 (0.88, 1.37)
	MOCEH	**0.87 (0.76, 1.00)**	ref	0.96 (0.81,1.14)	1.13 (0.91, 1.40)
	Pooled OR	-	ref	0.95(0.81,1.13)	1.13(0.92,1.40)
	Excluded cohort	*Outcome: High birth weight (Odds Ratio)*			
	DNBC	0.72 (0.46,1.11)	ref	0.84 (0.67,1.06)	0.83 (0.65,1.06)
	ABCD	1.03 (0.98, 1.09)	ref	1.01 (0.93,1.09)	0.88 (0.59, 1.32)
	INMA	0.95 (0.73, 1.25)	ref	0.97 (0.85, 1.12)	0.85 (0.62, 1,16)
	MOCEH	0.90 (0.65, 1.25)	ref	0.94 (0.80, 1.11)	0.99 (0.89, 1.09)
	Pooled OR	0.92(0.68,1.24)	ref	0.96(0.83,1.10)	0.93(0.78,1.11)
Pregnancy duration	Excluded cohort	*Outcome : Gestational age at birth (Hazard ratio)*			
	DNBC	1.02 (0.89,1.17)	ref	1.02 (0.94,1.12)	1.02 (0.94,1.09)
	ABCD	0.99 (0.97,1.01)	ref	1.02 (0.97, 1.09)	1.03 (0.99, 1.07)
	INMA	0.99 (0.97 1.01)	ref	**1.05 (1.02, 1.08)**	1.02 (0.98, 1.06)
	MOCEH	0.99 (0.97, 1.01)	ref	**1.04 (1.00, 1.07)**	1.02 (0.98, 1.06)
	Pooled HR	0.99 (0.97,1.01)	ref	**1.04 (1.01,1.07)**	1.02 (0.98,1.06)
	Excluded cohort	*Outcome: Preterm birth(Odds Ratio)*			
	DNBC	0.96 (0.47,1.98)	ref	1.04 (0.75,1.45)	1.43 (0.76,2.68)
	ABCD	0.96 (0.86, 1.08)	ref	1.14 (0.99, 1.31)	1.49 (0.95, 2.33)
	INMA	0.96 (0.86,1.07)	ref	1.11 (0.97, 1.28)	1.12(0.93, 1.35)
	MOCEH	0.96 (0.86, 1.07)	ref	1.11 (0.97, 1.28)	1.20 (0.78, 1.87)
	Pooled OR	0.96(0.86,1.07)	ref	1.12(0.97,1.28)	1.28(0.87,1.88)
	Excluded cohort	*Outcome: Postterm birth(Odds Ratio)*			
	DNBC	-	ref	0.94 (0.51,1.74)	1.17 (0.71,1.95)
	ABCD	-	ref	0.97 (0.67, 1.40)	1.12 (0.71, 1.77)
	INMA	-	ref	**0.84 (0.73, 0.95)**	0.95 (0.80, 1.13)
	MOCEH	-	ref	**0.85 (0.75, 0.97)**	0.98 (0.83, 1.16)
	Pooled OR	-	ref	**0.85(0.75,0.97)**	0.98(0.83,1.16)

	Exposure	None	Low	Intermediate	High
Fetal growth	Excluded cohort	*Outcome: Birth weight ratio (Mean difference)*			
	DNBC	-0.00 (-0.02,0.01)	ref	-0.00 (-0.01, 0.01)	0.00 (-0.01,0.01)
	ABCD	0.00 (-0.00, 0.00)	ref	-0.00 (-0.01 ,0.00)	0.00 (-0.00, 0.01)
	INMA	0.00 (-0.00,0.00)	ref	-0,00 (-0.00, 0.00)	-0.00 (-0.00, 0.00)
	MOCEH	0.00 (-0.00, 0.00)	ref	-0.00 (-0.00, 0.00)	0.00 (-0.00, 0.01)
	Pooled MD	0.00(-0.00,0.00)	ref	-0.00(-0.00,0.00)	0.00(-0.00,0.00)
	Excluded cohort	*Outcome: Small for gestational age (Odds Ratio)*			
	DNBC	0.80 (0.47,1.37)	ref	1.12 (0.84,1.48)	1.08 (0.83,1.40)
	ABCD	0.94 (0.86, 1.03)	ref	1.06 (0.82,1.36)	0.94 (0.74, 1.19)
	INMA	0.94 (0.86,1.03)	ref	0.97 (0.86, 1.09)	0.90 (0.74, 1.10)
	MOCEH	0.94 (0.86, 1.03)	ref	1.07 (0.87, 1.32)	0.90 (0.75, 1.08)
	Pooled OR	0.94(0.86,1.03)	ref	1.03(0.88,1.21)	0.94(0.78,1.13)
	Excluded cohort	*Outcome: Large for gestational age (Odds Ratio)*			
	DNBC	0.84 (0.53,1.31)	ref	0.93 (0.75,1.15)	0.96 (0.76, 1.21)
	ABCD	0.98 (0.92, 1.05)	ref	0.97 (0.89, 1.06)	0.93 (0.83, 1.05)
	INMA	0.98 (0.92, 1.04)	ref	0.97 (0.89, 1.06)	0.91 (0.81, 1.03)
	MOCEH	0.98 (0.92, 1.04)	ref	0.97 (0.89, 1.05)	0.93 (0.83, 1.04)
	Pooled OR	0.98(0.92,1.04)	ref	0.97(0.89,1.05)	0.93(0.83,1.04)

Abbreviations: HR, hazard ratio; MD, mean difference; OR, odds ratio; ref, reference level
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height
Statistically significant results marked with bold

Table S10. Meta-analysis of adjusted estimates -stratified by timing of cell phone use data collection, for the effect of maternal mobile phone use during pregnancy on birth outcomes.

	Exposure	None	Low	Intermediate	High
Birth weight	*Outcome: Birth weight within non-preterm neonates (Mean Difference in grams)*				
	Retrospective (DNBC & ABCD)	-20.76 (-87.35, 45.84)	ref	-7.53 (-21.78,6.72)	-2.15 (-21.46,17.15)
	Prospective (INMA & MOCEH)	1.21 (-100.70,103.12)	ref	-10.17 (-52.40,32.07)	-4.28 (-43.74,35.19)
	Pooled MD	-11.15(-53.24,30.94)	ref	-8.17(-21.34,5.00)	-2.56(-19.90,14.78)
	Outcome: Low birth weight (Odds Ratio)				
	Retrospective (DNBC & ABCD)	**0.87 (0.76,1.00)**	ref	0.96 (0.80,1.15)	1.09 (0.87,1.37)
	Prospective (INMA & MOCEH)	-	ref	0.92 (0.59,1.45)	1.36 (0.82,2.25)
	Pooled OR	-	ref	0.95(0.81,1.13)	1.13(0.92,1.40)
	Outcome: High birth weight (Odds Ratio)				
	Retrospective (DNBC & ABCD)	0.95 (0.71,1.26)	ref	0.96 (0.81,1.14)	0.96 (0.82,1.13)
	Prospective (INMA & MOCEH)	0.53 (0.12,2.36)	ref	0.87 (0.59,1.28)	0.67 (0.24,1.85)
	Pooled OR	0.92(0.68,1.24)	ref	0.96 (0.83,1.10)	0.93(0.78,1.11)

	Exposure	None	Low	Intermediate	High
Pregnancy duration	*Outcome : Gestational age at birth (Hazard ratio)*				
	Retrospective (DNBC & ABCD)	0.99 (0.97,1.01)	ref	**1.04 (1.01,1.08)**	1.02 (0.98,1.06)
	Prospective (INMA & MOCEH)	0.92 (0.71,1.19)	ref	1.00 (0.88,1.13)	1.05 (0.94,1.16)
	Pooled HR	0.99 (0.97,1.01)	ref	**1.04 (1.01,1.07)**	1.02 (0.98,1.06)
	Outcome: Preterm birth(Odds Ratio)				
	Retrospective (DNBC & ABCD)	0.96 (0.86, 1.07)	ref	1.11 (0.96,1.28)	1.06 (0.82,1.38)
	Prospective (INMA & MOCEH)	1.58 (0.45,5.61)	ref	1.21 (0.74, 1.97)	**2.03 (1.22,3.39)**
	Pooled OR	0.96(0.86,1.07)	ref	1.12(0.97,1.28)	1.28(0.87,1.88)
	Outcome: Postterm birth(Odds Ratio)				
	Retrospective (DNBC & ABCD)	0.98 (0.89, 1.07)	ref	**0.83 (0.73,0.95)**	0.95 (0.80,1.12)
	Prospective (INMA & MOCEH)	0.00 (0.00,Inf)	ref	1.32 (0.75,2.32)	1.60 (0.86,2.97)
	Pooled OR	-	ref	**0.85(0.75,0.97)**	0.98(0.83,1.16)
Fetal growth	*Outcome: Birth weight ratio (Mean difference)*				
	Retrospective (DNBC & ABCD)	0.00 (-0.00,0.00)	ref	-0.00 (-0.00,0.00)	0.00 (-0.00,0.01)
	Prospective (INMA & MOCEH)	0.01(-0.02, 0.03)	ref	-0.00 (-0.01,0.01)	0.00 (-0.01,0.01)
	Pooled MD	0.00(-0.00,0.00)	ref	-0.00(-0.00,0.00)	0.00(-0.00,0.00)
	Outcome: Small for gestational age (Odds Ratio)				
	Retrospective (DNBC & ABCD)	0.94 (0.86, 1.03)	ref	0.98 (0.87,1.10)	0.86 (0.73, 1.02)
	Prospective (INMA & MOCEH)	0.63 (0.19, 2.11)	ref	1.15 (0.72, 1.84)	1.14 (0.80,1.62)
	Pooled OR	0.94(0.86,1.03)	ref	1.03(0.88,1.21)	0.94(0.78,1.13)
	Outcome: Large for gestational age (Odds Ratio)				
	Retrospective (DNBC & ABCD)	0.98 (0.92,1.04)	ref	0.97 (0.89,1.06)	0.91 (0.81,1.03)
	Prospective (INMA & MOCEH)	0.93 (0.42,2.07)	ref	0.93 (0.70,1.24)	1.04 (0.75,1.43)
	Pooled OR	0.98(0.92,1.04)	ref	0.97(0.89,1.05)	0.93(0.83,1.04)

Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: HR, hazard ratio; Inf, infinity; MD, mean difference; OR, odds ratio; ref, reference level. Statistically significant results marked with bold

Table S11. Meta-analysis of adjusted estimates for the effect of cell mobile phone use during pregnancy on birth outcomes, using binary exposure variable.

	Exposure	Low	High		Exposure	Low	High		Exposure	Low	High	
	Outcome: Birth weight within non-preterm neonates (Mean Difference in grams)			P-value	*Outcome: Low birth weight (Odds Ratio)*			P-value	*Outcome: High birth weight (Odds Ratio)*			P-value
Birth weight	DNBC	ref	-10.25 (-21.60, 1.10)	0,08	DNBC	ref	1.11 (0.96,1.28)	0,15	DNBC	ref	0.99 (0.93,1.05)	0,69
	ABCD	ref	7.67 (-29.74,45.07)	0,69	ABCD	ref	0.92 (0.57,1.49)	0,74	ABCD	ref	0.86 (0.69,1.08)	0,20
	INMA	ref	-15.14 (-54.15, 23.87)	0,45	INMA	ref	1.14 (0.72,1.80)	0,58	INMA	ref	0.94 (0.62,1.43)	0,78
	MOCEH	ref	0.52 (-51.69,52.73)	0,98	MOCEH	ref	0.90 (0.37,2.17)	0,82	MOCEH	ref	0.74 (0.38,1.44)	0,39
	Pooled MD	ref	-8.83 (-19.08,1.43)	0,09	Pooled OR	ref	1.09 (0.96,1.24)	0,17	Pooled OR	ref	0.96 (0.89,1.05)	0,41

	Exposure	Low	High		Exposure	Low	High		Exposure	Low	High	
	Outcome : Gestational age at birth (Hazard ratio)			P-value	*Outcome: Preterm birth (Odds Ratio)*			P-value	*Outcome: Postterm birth (Odds Ratio)*			P-value
Pregnancy duration	DNBC	ref	**1.05 (1.02,1.07)**	**<0.01**	DNBC	ref	**1.17 (1.04,1.31)**	**<0.01**	DNBC	ref	**0.89 (0.81,0.99)**	**0,03**
	ABCD	ref	1.02 (0.94,1.11)	0,60	ABCD	ref	0.90 (0.61,1.33)	0,60	ABCD	ref	0.77 (0.51,1.16)	0,21
	INMA	ref	0.97 (0.88,1.07)	0,59	INMA	ref	1.43 (0.84,2.46)	0,19	INMA	ref	1.52 (0.89,2.59)	0,13
	MOCEH	ref	1.10 (0.95,1.27)	0,21	MOCEH	ref	1.44 (0.71,2.91)	0,31	MOCEH	ref	1.39 (0.14,13.75)	0,78
	Pooled HR	ref	**1.04 (1.02,1.07)**	**<0.01**	Pooled OR	ref	**1.16 (1.05,1.29)**	**<0.01**	Pooled OR	ref	0.95 (0.73, 1.23)	0,69

	Exposure	Low	High		Exposure	Low	High		Exposure	Low	High	
	Outcome: Birth weight ratio (Mean difference)			P-value	*Outcome: SGA (Odds Ratio)*			P-value	*Outcome: LGA (Odds Ratio)*			P-value
Fetal growth	DNBC	ref	-0.00 (-0.00,0.00)	0,31	DNBC	ref	0.97 (0.88,1.07)	0,58	DNBC	ref	0.97 (0.90,1.04)	0,38
	ABCD	ref	0.00 (-0.01,0.01)	0,67	ABCD	ref	1.03 (0.76,1.41)	0,84	ABCD	ref	0.93 (0.71,1.22)	0,60
	INMA	ref	-0.00 (-0.01,0.01)	0,67	INMA	ref	1.34 (0.95,1.87)	0,09	INMA	ref	1.01 (0.74,1.38)	0,97
	MOCEH	ref	0.00 (-0.02,0.02)	0,94	MOCEH	ref	0.99 (0.61,1.63)	0,95	MOCEH	ref	0.92 (0.59,1.45)	0,73
	Pooled MD	ref	-0.00 (-0.00,0.00)	0,37	Pooled OR	ref	1.03 (0.90,1.17)	0,70	Pooled OR	ref	0.97 (0.91,1.03)	0,32

Abbreviations: HR, hazard ratio; MD, mean difference; OR, odds ratio; ref, reference level
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height
Statistically significant results marked with bold

Table S12. Results of adjusted analysis for the effect of maternal cell phone use during pregnancy on fetal growth, based on observed birth weight per gestational age percentiles in DNBC.

	Exposure	None	Low	Intermediate	High	
	Outcome: Birth weight ratio (Mean difference)					P-value
Fetal growth	DNBC (observed percentiles)	0.00 (-0.00,0.00)	reference	-0.00 (-0.01,0.00)	0.00 (-0.01,0.01)	0,83
	Outcome: Small for gestational age (Odds Ratio)					P-value
	DNBC (observed percentiles)	0.96 (0.89,1.03)	reference	1.04 (0.93,1.15)	0.88 (0.76,1.02)	0,14
	Outcome: Large for gestational age (Odds Ratio)					P-value
	DNBC (observed percentiles)	0.99 (0.92,1.07)	reference	1.06 (0.95,1.18)	0.93 (0.79, 1.09)	0,42

Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height

5

Table S13. Meta-analysis of adjusted estimates for the effect of cell phone use during pregnancy on the odds of giving birth to a low birth weight neonate; analysis restricted to non-preterm neonates

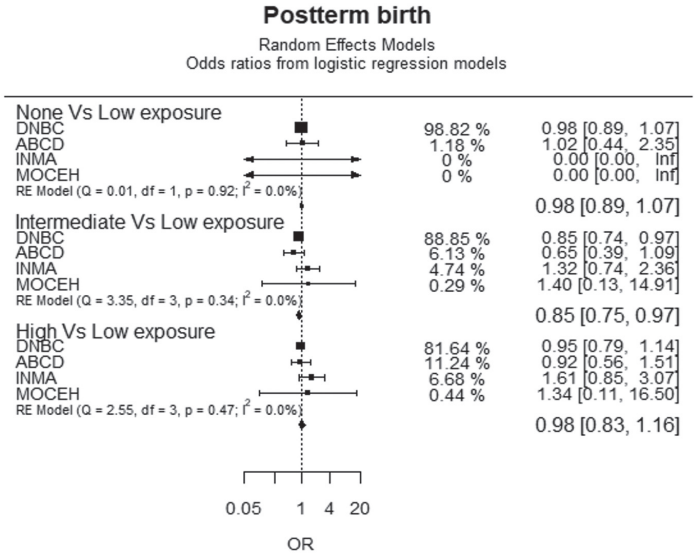
Exposure	None	Low	Intermediate	High	
	Outcome: Low birth weight (Odds Ratio)				p for trend
DNBC	1.00(0.78,1.28)	ref	0.94(0.67,1.33)	1.13(0.72,1.76)	0,87
ABCD	1.47(0.38,5.62)	ref	0.97(0.37,2.54)	1.10(0.42,2.88)	0,91
INMA	1.94(0.42,8.97)	ref	1.09(0.59,2.03)	0.90(0.40,2.01)	0,61
MOCEH	0.00(0.00,Inf)	ref	0.43(0.08,2.31)	0.76(0.13,4.58)	0,84
Pooled OR	1.03 (0.81,1.31) [a]	ref	0.95(0.72,1.27)	1.06(0.74,1.51)	0,98

[a] excluding MOCEH
Abbreviations: Inf, infinity; OR, odds ratio; ref, reference level
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height

Figure S1. Meta-analysis results for the effect of cell mobile phone use during pregnancy on gestational age at birth.

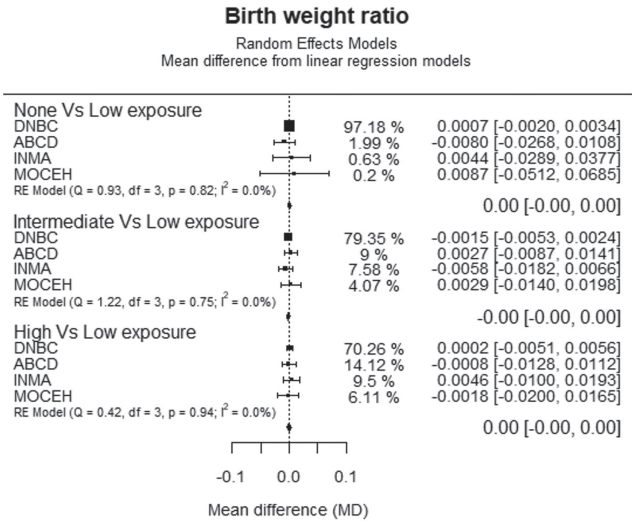
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; HR, hazard ratio; BMI, Body Mass Index

Figure S2. Meta-analysis results for the effect of cell mobile phone use during pregnancy on the odds of postterm birth.



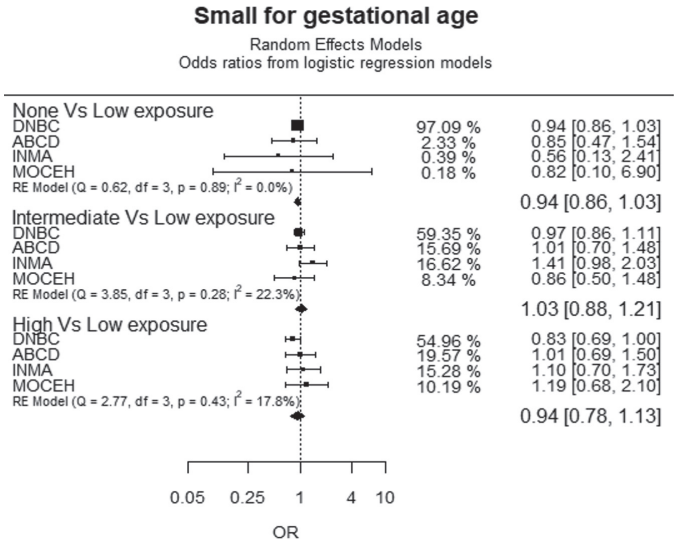
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; OR, odds ratio; BMI, Body Mass Index

Figure S3. Meta-analysis results for the effect of cell mobile phone use during pregnancy on birth weight ratio.



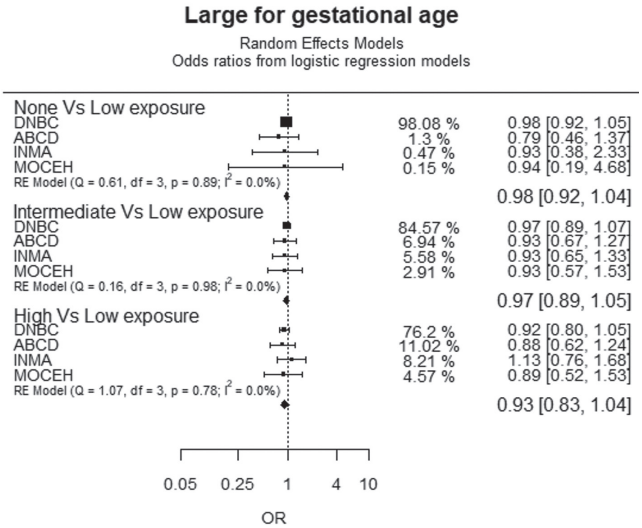
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; BMI, Body Mass Index

Figure S4. Meta-analysis results for the effect of maternal cell phone use during pregnancy on the odds of giving birth to a small for gestational age neonate.



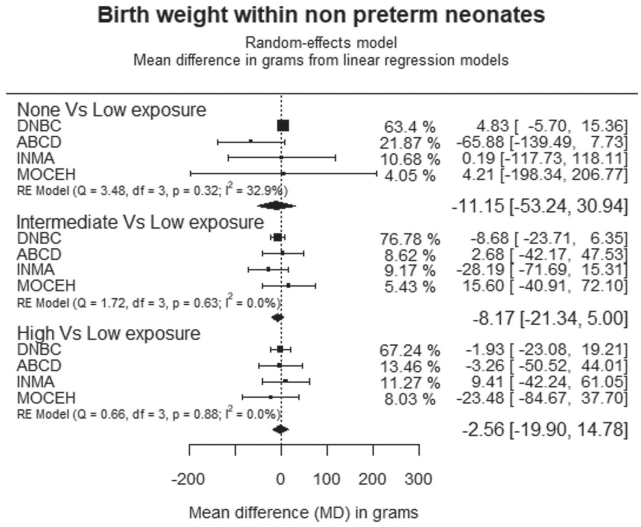
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; OR, odds ratio; BMI, Body Mass Index

Figure S5. Meta-analysis results for the effect of maternal cell phone use during pregnancy on the odds of giving birth to a large for gestational age neonate.



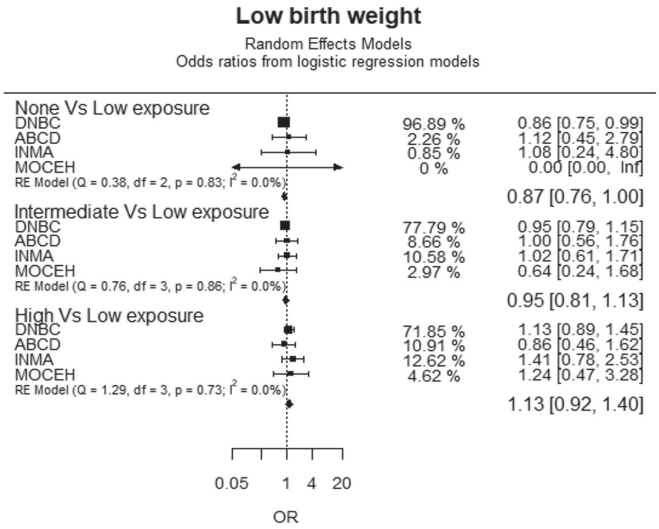
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; OR, odds ratio; BMI, Body Mass Index

Figure S6. Meta-analysis results for the effect of maternal cell phone use during pregnancy on birth weight; analyses restricted to non preterm neonates.



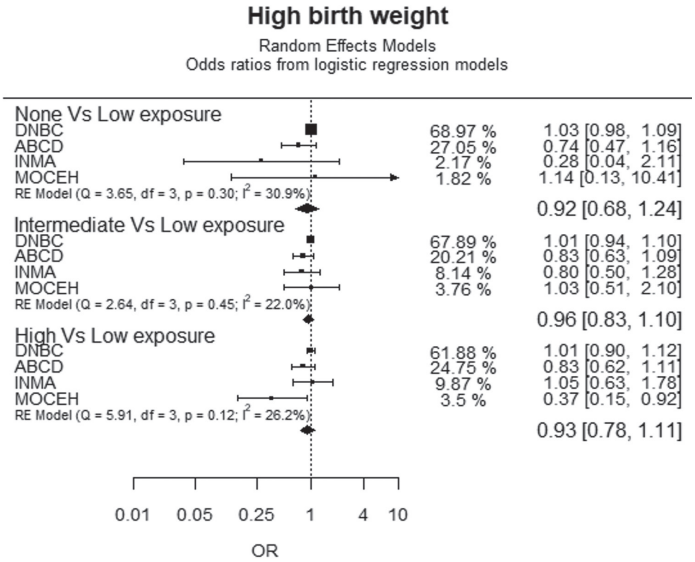
Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; BMI, Body Mass Index

Figure S7. Meta-analysis results for the effect of maternal cell phone use during pregnancy on the odds of giving birth to a low birth weight neonate.



Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; OR, odds ratio; BMI, Body Mass Index

Figure S8. Meta-analysis results for the effect of maternal cell phone use during pregnancy on the odds of giving birth to a high birth weight neonate.



Results adjusted for maternal age, parity, active and passive smoking, alcohol consumption, pre-pregnancy BMI, educational level and socioeconomic position, marital status, and maternal height. Abbreviations: RE, random effects; df, degrees of freedom; OR, odds ratio; BMI, Body Mass Index

REFERENCES

1. World Bank. Mobile cellular subscriptions (per 100 people). 2016;(https://data.worldbank.org/indicator/IT.CEL.SETS.P2). (Accessed February 9, 2018)
2. Health Protection Agency. Health Effects from Radiofrequency Electromagnetic Fields. 2012 1-348 p.(http://www.hpa.org.uk/webc/hpawebfile/hpaweb_c/1317133827077)
3. World Health Organization. Electromagnetic fields and public health: mobile phones. 2014;(http://www.who.int/mediacentre/factsheets/fs193/en/)
4. Lauer O, Frei P, Gosselin MC, et al. Combining near- and far-field exposure for an organ-specific and whole-body RF-EMF proxy for epidemiological research: A reference case. *Bioelectromagnetics*. 2013;34(5):366–374.
5. Daşda S, Akda MZ, Ayyildiz O, et al. Do cellular phones alter blood parameters and birth weight of rats? *Electro- and Magnetobiology*. 2000;
6. Yüksel M, Nazıroğlu M, Özkaya MO. Long-term exposure to electromagnetic radiation from mobile phones and Wi-Fi devices decreases plasma prolactin, progesterone, and estrogen levels but increases uterine oxidative stress in pregnant rats and their offspring. *Endocrine*. 2016;
7. Shirai T, Wang J, Kawabe M, et al. No adverse effects detected for simultaneous whole-body exposure to multiple-frequency radiofrequency electromagnetic fields for rats in the intrauterine and pre- and post-weaning periods. *J. Radiat. Res.* 2017;
8. Sommer AM, Grote K, Reinhardt T, et al. Effects of Radiofrequency Electromagnetic Fields (UMTS) on Reproduction and Development of Mice: A Multi-generation Study. *Radiat. Res.* 2009;
9. Cabot E, Christ A, Bühlmann B, et al. Quantification Of RF-exposure of the Fetus Using Anatomical CAD-Models in Three Different Gestational Stages. *Health Phys.* [electronic article]. 2014;107(5):369–381. (http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00004032-201411000-00001). (Accessed February 9, 2018)
10. Nagaoka T, Togashi T, Saito K, et al. An anatomically realistic whole-body pregnant-woman model and specific absorption rates for pregnant-woman exposure to electromagnetic plane waves from 10 MHz to 2 GHz. *Phys. Med. Biol.* 2007;
11. Tateno A, Tanaka K, Nagaoka T, et al. Specific absorption rates of pregnant females and their fetuses from simple and realistic electromagnetic sources. *IEICE Commun. Express*. 2014;3(2):55–60.
12. Varsier N, Dahdouh S, Serrurier A, et al. Influence of pregnancy stage and fetus position on the whole-body and local exposure of the fetus to RF-EMF. *Phys. Med. Biol.* [electronic article]. 2014;59(17):4913–4926. (http://www.ncbi.nlm.nih.gov/pubmed/25098501). (Accessed January 23, 2018)
13. Luo Q, Jiang Y, Jin M, et al. Proteomic analysis on the alteration of protein expression in the early-stage placental villous tissue of electromagnetic fields associated with cell phone exposure. *Reprod. Sci.* [electronic article]. 2013;20(9):1055–1061. (http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L369611216). (Accessed January 25, 2018)
14. Emekaya MA, Seyhan N, Ömeroğlu S. Pulse modulated 900 MHz radiation induces hypothyroidism and apoptosis in thyroid cells: A light, electron microscopy and immunohistochemical study. *Int. J. Radiat. Biol.* 2010;86(12):1106–1116.
15. Misa-Agustiño MJ, Jorge-Mora T, Jorge-Barreiro FJ, et al. Exposure to non-ionizing radiation provokes changes in rat thyroid morphology and expression of HSP-90. *Exp. Biol. Med.* 2015;240(9):1123–1135.
16. Nilgun C-A. Evaluation of factors affecting birth weight and preterm birth in southern Turkey. *JPMA* [electronic article]. 2013;63(4). (http://jpma.org.pk/PdfDownload/4113.pdf). (Accessed January 21, 2018)

17. Mortazavi SMJ, Shirazi KR, Mortazavi G. The study of the effects of ionizing and non-ionizing radiations on birth weight of newborns to exposed mothers. *J. Nat. Sci. Biol. Med.* [electronic article]. 2013;4(1):213–217. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3633280&tool=pmcentrez&rendertype=abstract)

18. Baste V, Oftedal G, Mollerlokken O, et al. Prospective Study of Pregnancy Outcomes after Parental Cell Phone Exposure: The Norwegian Mother and Child Cohort Study. *Epidemiology*. 2015;26(4):613–621.

19. "Generalized EMF research using novel methods". GERoNiMO. 2014;(http://radiation.isglobal.org/index.php/geronimo-home). (Accessed February 15, 2018)

20. Olsen J, Melbye M, Olsen SF, et al. The Danish National Birth Cohort ± its background, structure and aim. *Scand J Public Heal. Taylor Fr.* 2001;29:301–7311.

21. Divan HA, Kheifets L, Obel C, et al. Prenatal and postnatal exposure to cell phone use and behavioral problems in children. *Epidemiology*. 2008;19(4):523–529.

22. Van Eijsden M, Vrijkotte TGM, Gemke RJBJ, et al. Cohort profile: The Amsterdam born children and their development (ABCD) study. *Int. J. Epidemiol.* 2011;

23. Guxens M, Ballester F, Espada M, et al. Cohort profile: The INMA-INfancia y Medio Ambiente-(environment and childhood) project. *Int. J. Epidemiol.* 2012;

24. Kim BM, Ha M, Park HS, et al. The Mothers and Children's Environmental Health (MOCEH) study. *Eur. J. Epidemiol.* 2009;

25. Visser GHA, Eilers PHC, Elferink-Stinkens PM, et al. New Dutch reference curves for birthweight by gestational age. *Early Hum. Dev.* [electronic article]. 2009;85(12):737–44. (http://linkinghub.elsevier.com/retrieve/pii/S0378378209001960). (Accessed February 9, 2018)

26. Carrascosa Lezcano A, Ferrández Longás A, Yeste Fernández D, et al. Estudio transversal español de crecimiento 2008. Parte I: Valores de peso y longitud en recién nacidos de 26-42 semanas de edad gestacional. *An. Pediatr.* 2008;68(6):544–551.

27. Lim JS, Lim SW, Ahn JH, et al. New Korean reference for birth weight by gestational age and sex: data from the Korean Statistical Information Service (2008-2012). *Ann. Pediatr. Endocrinol. Metab.* [electronic article]. 2014;19(3):146–53. (http://www.ncbi.nlm.nih.gov/pubmed/25346919%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4208258)

28. Skjærven R, Gjessing' HK, Bakketeig LS. Birthweight by gestational age in Norway. *Acta Obstet. Gynecol. Scand.* 2000;79(6):440–449.

29. Birks L, Guxens M, Papadopoulou E, et al. Maternal cell phone use during pregnancy and child behavioral problems in five birth cohorts. *Environ. Int.* 2017;

30. Voskamp BJ, Kazemier BM, Schuit E, et al. Birth weight ratio as an alternative to birth weight percentile to express infant weight in research and clinical practice: a nationwide cohort study. *Obstet. Gynecol. Int.* [electronic article]. 2014;2014:749476. (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4147261&tool=pmcentrez&rendertype=abstract)

31. Barba C, Cavalli-Sforza T, Cutter J, et al. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet*. 2004;363(9403):157–163.

32. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev. Sci.* 2007;8(3):206–213.

33. Li K-H, Meng X-L, Raghunathan TE, et al. Significance levels from repeated p-values with multiply-imputed data. *Stat. Sin.* 1991;1:65–92. (http://www.jstor.org/stable/24303994). (Accessed May 28, 2018)

34. Joseph KS, Kramer MS. The fetuses-at-risk approach: Survival analysis from a fetal perspective. *Acta Obstet. Gynecol. Scand.* 2017;1–12.

35. R Core Team. R: A Language and Environment for Statistical Computing. 2017;(https://www.r-project.org/)

36. Yoshida K, Bohn. J. Create "Table 1" to Describe Baseline Characteristics. 2015;(https://cran.r-project.org/package=tableone)

37. van Buuren S, Groothuis-Oudshoorn K. Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* [electronic article]. 2011;45(3):1–67. (http://www.jstatsoft.org/v45/i03/)

38. Robitzsch A, Grund S, Henke T. miceadds: Some Additional Multiple Imputation Functions, Especially for "mice." 2017;(https://cran.r-project.org/package=miceadds)

39. Therneau TM. A Package for Survival Analysis in S. 2015;(https://cran.r-project.org/package=survival)

40. Terry M. Therneau, Patricia M. Grambsch. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.

41. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* [electronic article]. 2010;36(3):1–48. (http://www.jstatsoft.org/v36/i03/)

42. Bolte JFB, Eikelboom T. Personal radiofrequency electromagnetic field measurements in the Netherlands: Exposure level and variability for everyday activities, times of day and types of area. *Environ. Int.* 2012;

43. De Graaf JP, Steegers EAP, Bonsel GJ. Inequalities in perinatal and maternal health. *Curr. Opin. Obstet. Gynecol.* 2013;

44. Lu X, Oda M, Ohba T, et al. Association of excessive mobile phone use during pregnancy with birth weight: An adjunct study in Kumamoto of Japan Environment and Children's Study. *Environ. Health Prev. Med.* [electronic article]. 2017;22(1). (https://www-ncbi-nlm-nih-gov.proxy.library.uu.nl/pmc/articles/PMC5664573/pdf/12199_2017_Article_656.pdf). (Accessed January 21, 2018)

45. Morgen CS, Bjork C, Andersen PK, et al. Socioeconomic position and the risk of preterm birth--a study within the Danish National Birth Cohort. *Int. J. Epidemiol.* 2008;

46. Chen S, Zhou X, Zhu H, et al. Preconception TSH and pregnancy outcomes: a population-based cohort study in 184 611 women. *Clin. Endocrinol. (Oxf).* [electronic article]. 2017;86(6):816–824. (http://www.ncbi.nlm.nih.gov/pubmed/28295470). (Accessed February 27, 2018)

47. Arbib N, Hadar E, Sneh-Arbib O, et al. First trimester thyroid stimulating hormone as an independent risk factor for adverse pregnancy outcome. *J. Matern. Neonatal Med.* [electronic article]. 2017;30(18):2174–2178. (http://www.ncbi.nlm.nih.gov/pubmed/27677438). (Accessed February 27, 2018)

48. Schneuer FJ, Nassar N, Tasevski V, et al. Association and Predictive Accuracy of High TSH Serum Levels in First Trimester and Adverse Pregnancy Outcomes. *J. Clin. Endocrinol. Metab.* [electronic article]. 2012;97(9):3115–3122. (http://www.ncbi.nlm.nih.gov/pubmed/22723328). (Accessed February 27, 2018)

49. Casey BM, Dashe JS, Wells CE, et al. Subclinical Hypothyroidism and Pregnancy Outcomes. *Obstet. Gynecol.* [electronic article]. 2005;105(2):239–245. (http://www.ncbi.nlm.nih.gov/pubmed/15684146). (Accessed February 27, 2018)

50. SCENIHR (Scientific Committee on Emerging and Newly Identified Healtlh Risks). Potential health effects of exposure to electromagnetic fields (EMF). 2015;(http://ec.europa.eu/health/scientific_committees/policy/index_en.htm). (Accessed April 3, 2018)

51. Thomée S, Härenstam A, Hagberg M. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults - A prospective cohort study. *BMC Public Health* [electronic article]. 2011;11(1):66. (http://www.biomedcentral.com/1471-2458/11/66)

52. Medsker B, Forno E, Simhan H, et al. Prenatal Stress, Prematurity, and Asthma. *Obstet. Gynecol. Surv.* 2015;70(12):773–779. (http://www.ncbi.nlm.nih.gov/pubmed/26676148). (Accessed March 1, 2018)

53. Beydoun H, Saftlas AF. Physical and mental health outcomes of prenatal maternal stress in human and animal studies: A review of recent evidence. *Paediatr. Perinat. Epidemiol.* 2008;22(5):438–466.

54. Mulder E, Robles de Medina P, Huizink A, et al. Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Hum. Dev.* 2002;70(1–2):3–14.

55. Kristenson M, Eriksen HR, Sluiter JK, et al. Psychobiological mechanisms of socioeconomic differences in health. *Soc. Sci. Med.* 2004;58(8):1511–1522.

56. Lobel M, Cannella DL, Graham JE, et al. Pregnancy-Specific Stress, Prenatal Health Behaviors, and Birth Outcomes. *Heal. Psychol.* 2008;27(5):604–615.

57. Lockwood C, Ramin S, Barss V. Stress-associated preterm delivery: the role of corticotropin-releasing hormone. *Am J Obs. Gynecol*. 1999;180(1):S264–S266.

58. Wadhwa PD, Entringer S, Buss C, et al. The contribution of maternal stress to preterm birth: Issues and considerations. *Clin. Perinatol.* 2011;38(3):351–384.

59. Choi K-H, Ha M, Burm E, et al. Multiple assessment methods of prenatal exposure to radio frequency radiation from telecommunication in the Mothers and Children's Environmental Health (MOCEH) study. *Int. J. Occup. Med. Environ. Health* [electronic article]. 2016;29(6):959–972. (http://www.ncbi.nlm.nih.gov/pubmed/27869246). (Accessed January 23, 2018)

60. Vrijheid M, Cardis E, Armstrong BK, et al. Validation of short term recall of mobile phone use for the Interphone study. *Occup. Environ. Med.* 2006;63(4):237–243.

61. Goedhart G, Kromhout H, Wiart J, et al. Validating self-reported mobile phone use in adults using a newly developed smartphone application. *Occup. Environ. Med.* 2015;72(11):812–818.

62. Goedhart G, van Wel L, Langer CE, et al. Recall of mobile phone usage and laterality in young people: The multinational Mobi-Expo study. *Environ. Res.* [electronic article]. 2018;165(November 2017):150–157. (https://doi.org/10.1016/j.envres.2018.04.018)

Chapter 6

Modeled and perceived RF-EMF, noise and air pollution and symptoms in a population cohort. *Is perception key in predicting symptoms?*

Marije Reedijk[a], Astrid L. Martens[a], Tjabe Smid, Anke Huss, Danielle Timmermans, Maciej Strak, Wim Swart, Virissa Lenters, Hans Kromhout, Robert Verheij, Pauline Slottje[b], Roel C.H. Vermeulen[b]

[a] Equal contribution
[b] Equal contribution

Published: Science of the total environment, 2018, vol. 639: 75-83.

ABSTRACT

Psychosocial research has shown that perceived exposure can influence symptom reporting, regardless of actual exposure. The impact of this phenomenon on the interpretation of results from epidemiological research on environmental determinants of symptoms is unclear. Our aim was to compare associations between modeled exposures, the perceived level of these exposures and reported symptoms (non-specific symptoms, sleep disturbances, and respiratory symptoms) for three different environmental exposures (radiofrequency electromagnetic fields (RF-EMF), noise, and air pollution). These environmental exposures vary in the degree to which they can be sensorially observed. Participant characteristics, perceived exposures, and self-reported health were assessed with a baseline (n=14,829, 2011/2012) and follow-up (n=7,905, 2015) questionnaire in the Dutch population-based Occupational and Environmental Health Cohort (AMIGO). Environmental exposures were estimated at the home address using spatial models. Cross-sectional and longitudinal regression models were used to examine the associations between modeled and perceived exposures, and reported symptoms. The extent to which exposure sources could be observed by participants likely influenced correlations between modeled and perceived exposure as correlations were moderate for air pollution (r_{Sp}=0.34) and noise (r_{Sp}=0.40), but less so for RF-EMF (r_{Sp}=0.11). Perceived exposures were consistently associated with increased symptom scores (respiratory, sleep, non-specific). Modeled exposures, except RF-EMF, were associated with increased symptom scores, but these associations disappeared or strongly diminished when accounted for perceived exposure in the analyses. Perceived exposure has an important role in symptom reporting. When environmental determinants of symptoms are studied without acknowledging the potential role of both modeled and perceived exposures, there is a risk of bias in health risk assessment. However, the etiological role of exposure perceptions in relation to symptom reporting requires further research.

Abbreviations

AMIGO, Occupational and Environmental Health Cohort Study; 4DSQ-S, somatization scale of the Four-Dimensional Symptom Questionnaire; ESCAPE, European Study of Cohorts for Air Pollution Effects; MOS, Medical Outcomes Study; STAMINA, Standard Model Instrumentation for Noise Assessments

INTRODUCTION

Radiofrequency electromagnetic fields (RF-EMF) from mobile phone base stations, noise exposure from road traffic, and air pollutants are environmental exposures often clustered in more densely populated area[1,2]. The general population is involuntarily exposed to these exposures, and many people have concerns about potential health risks. Recent studies have highlighted a complex interplay between these environmental exposures, perceptions of exposure and health risks, and symptom reporting[3–5]. For example, for residential RF-EMF exposure from mobile phone base stations we recently showed, using a longitudinal design, that perceived, but not modeled exposure, was associated with self-reported symptoms[5]. For noise from road traffic and air pollutants, perceptions mediated the effect of exposure on symptoms[3,4]. These studies show that research into environmental determinants of symptoms can benefit from applying insights from both psychosocial and epidemiological research disciplines.

The current study compares effects of RF-EMF from mobile phone base stations, noise and air pollutants from road traffic for the following symptom-based health outcomes: non-specific symptoms, sleep disturbance, and respiratory symptoms. These health outcomes are chosen based on variation in the plausibility of the link with the different environmental exposures. For environmental RF-EMF exposure, there is evidence of changes in sleep electroencephalography (EEG)[6], but no convincing epidemiological evidence for specific effects on symptoms, nor a known biological mechanism[7,8]. However, people who regard themselves as electrohypersensitive report a wide variety of non-specific symptoms, such as headache, fatigue, and pain which they attribute to EMF exposure[9,10]. Noise exposure on the other hand, can induce arousal, which can be observed during sleep through changes in EEG, heart rate, and respiration[11]. Prior epidemiological studies reported associations between noise exposure and sleep disturbances e.g.,[12–14], and there is also evidence for effects on wellbeing and overall symptoms[15]. Air pollutants can cause oxidative stress and an inflammatory response[16]. Epidemiological studies have found associations between exposure to air pollutants and respiratory symptoms such as shortness of breath, coughing, and wheezing[17–19].

The expectation that negative health effects may occur, can itself induce symptoms when people think they are exposed, regardless of the actual exposure and risk[20–22]. This is also described as nocebo-effect, as the counterpart of placebo[23]. Nocebo-effects may be part of a circular process, where experiencing symptoms can also influence perceptions of potential environmental health hazards[10,24]. Perceptions of environmental exposures, perceived health risks and worries play an important role in symptom experiences[25–27]. The type of symptoms that people report and associate with an environmental hazard differs depending on biological characteristics of the hazard and the content of media reports[22,28]. There are differences in the degree to which environmental exposures can be sensorially detected by humans, and this may affect perceived exposure. For RF-EMF from mobile phone base stations, only the exposure source can be perceived (f.i. visibility of antennas on nearby

buildings). While black smoke or diesel exhaust can sometimes be seen on windows, or smelt, there is no sensory system in humans that can directly perceive the level of air pollutants such as NO_2. Traffic noise is the only exposure, in this study, which is perceived by a specific sensory system in humans[14] and we therefore expect higher correlations with self-reported perceived exposure than for air pollutants and in particular RF-EMF.

Aims and research questions

This paper applies insights from epidemiological and psychosocial research to study environmental determinants of symptom-based health outcomes within the Dutch population-based Occupational and Environmental Health Cohort study (AMIGO). We have formulated three research questions, with the purpose of achieving a better understanding of the interplay between environmental exposures, perceptions and reported symptoms: 1) To what extent are participants able to assess personal exposure levels, and how does this differ between environmental exposures?; 2) What are the associations between modeled exposures and symptom-based health outcomes, and between perceived exposures and symptom-based health outcomes, and how do these associations change when both modeled and perceived exposures are taken into account simultaneously?; 3) Lastly, what is the impact on perceived exposures and on symptom-based health outcomes, after a change in exposures due to moving to a different home? With these final longitudinal analyses, we aim to improve our understanding of the processes that underlie the relations between modeled and perceived exposures, and symptom-based health outcomes.

MATERIAL AND METHODS

Study population

Data for this study were collected within the Dutch population-based AMIGO cohort. This cohort was set up in 2011 and 2012 to study environmental and occupational determinants of chronic diseases and symptoms in the general population (see[29] for a full description of the AMIGO cohort). Participants were recruited through general practices, and were 31-65 years old at baseline (T0, 2011/2012). Of the invited 93,849 people, 14,829 participants responded (participation rate=16%). A follow-up questionnaire was conducted in 2015 (T1, invited n=14,597, response=7,905; 54%), to assess changes in exposure perceptions and symptom-based health outcomes.

Symptom based-health outcomes

Self-reported symptom-based health outcomes (non-specific symptoms, sleep disturbances and respiratory symptoms) were assessed with the baseline and follow-up questionnaires. Non-specific symptoms were assessed with the somatization scale of the Four-Dimensional Symptom

Questionnaire (4DSQ-S)[30], which consists of 16 non-specific somatic symptoms commonly reported in general practices (e.g. headaches, low back pain, and dizziness). Participants indicated for each symptom whether they were bothered by it during the previous week on a 5-point scale. The scores per symptom were trichotomized (no= 0; sometimes= 1; regularly/often/constantly= 2) and then summed over the symptoms to obtain a total score. Sleep disturbances were measured with the items of the Medical Outcomes Study (MOS)[31]. The sleep problem index 9 was calculated following the instructions described in[31] as a measure of overall sleep quality. Higher scores indicate more sleep disturbance, or lower sleep quality. Respiratory symptoms were assessed with items from the European Community Respiratory Health Survey II[32]. A measure for respiratory symptoms is calculated as the sum of five items based on the method used by Sunyer et. al.[33]. A higher respiratory score indicates more respiratory symptoms.

Modeled environmental exposures

The home addresses were geocoded using data from the Netherlands Cadastre, Land Registry and Mapping Agency (Kadaster Netherlands). The geocoded home addresses were linked to various spatial models to assess exposure at the home addresses of participants as a proxy for actual exposure. Exposures were modeled for both baseline (2011/2012) and follow-up (2015) home addresses. For noise and air pollutants, the model estimates only changed if participants moved to a different home, as other input variables in the exposure models were not updated over time.

RF-EMF exposure from mobile phone base stations was modeled with the 3D-geospatial model NISMap. The applicability of this model for epidemiological studies has been described in a number of previous studies[35,36]. The model uses detailed information about 3D building data, topography, bedroom elevation, antenna location, antenna characteristics and radiation patterns to compute the field strength of GSM900 (Global System for Mobile Communication), GSM1800, and UMTS (Universal Mobile Telecommunications System) frequencies at the geocoded addresses in mW/m^2. Information about location and characteristics of antennas was available for 2011, 2012, and 2013. Input data closest to the questionnaire completion date was used for baseline and follow-up estimates. Road traffic noise exposure was estimated by the Standard Model Instrumentation for Noise Assessments (STAMINA), which is a model to map environmental noise from various sources in the Netherlands[37,38]. Input variables for the calculations were information on noise sources (for road traffic noise this includes traffic intensities, speed, composition and type of road surface), building data, and ground type. We used noise levels (dB) estimated over a whole day period (Lden), which uses penalties for the evening and night. In practice there is a very high correlation between whole day period noise estimates and night time noise estimates as shown in an earlier Dutch study (r_{Sp} of 0.99)[37]. Uncertainty in the modelling of noise at low levels and lack of information on roads with low volumes of traffic led to the introduction of a cut-off value of 24 dB Lden for the noise level.

Long-term residential ambient air pollutant concentrations of NO_2 (nitrogen dioxide), NO_X (total concentration of NO and NO_2), $PM_{2.5}$ and PM_{10} (particles with an aerodynamic diameter ≤ 2.5 µm and ≤ 10 µm, respectively) were assessed using land-use regression (LUR) models developed within the European Study of Cohorts for Air Pollution Effects (ESCAPE)[39,40], following a standardized protocol described elsewhere[39,40]. Air pollution measurements used to develop the LUR models took place between 2008 and 2011. NO_2 was the primary air pollutant metric, as this exposure is primarily traffic related, corresponding with our perceived exposure measure. Results for other air pollutants: NO_X, $PM_{2.5,}$ PM_{10} are reported in the Appendix.

Perceived environmental exposure

Perceived exposure was assessed at both time points (T0, 2011/2012 and T1, 2015) for the environmental exposures with the question: "To what extent are you exposed to: (1) electromagnetic fields/radiation from base stations for mobile phones, radio or television; (2) noise from road traffic in your home neighborhood; (3) air pollution in the residential area from road traffic?" Answers were given on a 7-point Likert scale ranging from 0= not at all, to 6= very much.

Covariates

The baseline questionnaire included questions on sex, age (in years), highest attained level of education (classification according to Statistics Netherlands), and smoking (never, ever, current). Low neighborhood income (percentage of income earners in the neighborhood with an income lower than the 40th percentile of the national income distribution) as an indication of neighborhood socioeconomic position was obtained from Statistics Netherlands[41].

Statistical analysis

We reported the baseline characteristics of the study participants and summary statistics for modeled and perceived environmental exposures (RF-EMF, noise, air pollutants), as well as the various health outcomes (non-specific symptoms, sleep disturbances, respiratory symptoms), for the two time points used in this study (T0, 2011/2012 and T1, 2015). We calculated Spearman correlations between all variables of interest at baseline (e.g. the correlation among all three modeled exposures, among all three perceived exposures, among the different symptom-based health outcomes, and the correlation between modeled and corresponding perceived exposure). The correlation between modeled exposure and the corresponding perceived exposure is interpreted as an indication of the accuracy in which participants were able to assess personal exposure levels.

Associations between modeled exposures, perceived exposures and the symptom-based health outcomes (second research question) were analyzed with mixed models. We performed both single predictor models (including modeled or perceived exposure, respectively) and two-predictor models (including modeled and perceived exposure simultaneously). We then used fixed effect models [42] in the follow-up sample to analyze temporal changes, i.e. whether intra-individual variation in perceived exposure was associated with intra-individual variation in health outcomes. Modeled exposure was not included in these analyses, as there was no temporal variation in estimates for air pollutants and noise unless participants moved to a different home address.

To assess the impact of a change in the environment on modeled, perceived exposures and symptom-based health outcomes (last research question), we analyzed only the group of participants who had moved between baseline and follow-up and participated in both questionnaires (n=592). Only for this group there were participants with sufficient temporal variation in modeled exposure estimates to evaluate the impact thereof on health outcomes. We first plotted the course of perceived exposure (means) over time for three percentile-based categories of absolute change (T1-T0) in modeled exposure (decrease: 0-20, no or small change: 20-80, increase: 80-100). Finally, we performed fixed effect models for the group of movers, including both modeled and perceived exposures as predictors.

Perceived and modeled exposures were analyzed as continuous variables with the exception of RF-EMF from mobile phone base stations. Because a large percentage of participants had modeled RF-EMF levels at or near 0.000 mW/m^2, values were dichotomized based on the 90th percentile of baseline exposure, with values $\leq$ 0.050 mW/m^2 defined as low and $>$ 0.050 mW/m^2 defined as high, similar to Martens et. al.[5]. The health outcomes are analyzed continuous. All mixed models were adjusted for sex, age, education, smoking, neighborhood income level, and for year of filling in the questionnaire (baseline/follow-up). The fixed effect model controls for all measured and unmeasured stable characteristics of an individual[42] and therefore no covariates were included in the model.

Missing values ranged between 0% and 7%. Missing values were imputed using the fully conditional method (FCS) in SAS. This method applied a discriminant function for categorical variables and predictive mean matching for continuous variables. A p-value of 0.05 was used as the cut-off for statistical significance. The statistical analyses were carried out using SAS (SAS Institute Inc., Cary, NC, USA).

RESULTS

Descriptive statistics

Baseline characteristics of the AMIGO participants at baseline (n=14,829) and follow-up (n=7,905) are shown in Table 1. Participants who filled in the follow-up questionnaire (follow-up sample) were more often higher educated, less often current smokers, on average older (Table 1), and had more favorable symptom scores at baseline (Table 2) than the baseline cohort. The follow-up sample had similar scores at baseline for modeled exposures and perceived exposures, compared to participants who participated only at baseline (Table 2). Over time, perceived exposures increased, and sleep disturbance and respiratory symptoms decreased in the follow-up sample. Modeled exposure values ranged from 0.00-3.13 mW/m^2 for RF-EMF, 27.00-74.80 dB for noise, and 10.25-68.39 µg/m^3 for NO_2 at baseline.

Table 1. General baseline (2011/2012) characteristics for the baseline cohort (n=14,829) and follow-up sample (n=7,905) in AMIGO.

	Baseline cohort (n=14829)		**Follow-up sample (n=7905)**	
Variable	*n*	*%*	*n*	*%*
Sex				
Male	6 561	44.24	3 728	47.16
Female	8 268	55.76	4 177	52.84
Education				
Low	4 714	31.79	2 246	28.41
Medium	4 773	32.19	2 420	30.61
High	5 342	36.02	3 239	40.97
Smoking status				
Never	6 748	45.51	3 685	46.62
Ever	5 755	38.81	3 239	40.97
Current smoker	2 326	15.69	981	12.41
	Mean (SD)	*IQR*	*Mean (SD)*	*IQR*
age (years)	50.65 (9.37)	43.00-59.00	52.17 (9.04)	46.00-60.00
socioeconomic position (%)[a]	39.41 (6.92)	35.00-44.00	39.16 (6.87)	34.00-44.00

[a] Percentage income-earners with a low-income in the neighborhood.

Table 2. Exposure and health outcome characteristics for the baseline AMIGO cohort (n=14,829) at T0 (2011/2012) and the follow-up sample (n=7,905) at T0 (2011/2012) and T1 (2015).

	Baseline cohort T0 (n=14,829)		Follow-up sample T0 (n=7,905)		Follow-up sample T1 (n=7,905)	
Variable	**Mean (SD)**	**IQR**	**Mean (SD)**	**IQR**	**Mean (SD)**	**IQR**
modeled RF-EMF (mW/m^2)	0.02 (0.09)	0.00-0.01	0.02 (0.09)	0.00-0.01	0.03 (0.11)	0.00-0.02
modeled Noise (dB)	53.11 (5.82)	49.40-56.70	53.15 (5.86)	49.40-56.70	53.14 (5.86)	49.40-56.70
modeled NO_2 (µg/m^3)	22.11 (5.60)	18.30-25.53	22.22 (5.59)	18.44-25.64	22.19 (5.60)	18.38-25.61
perceived Base station (0-6)	1.05 (1.26)	0.00-2.00	1.02 (1.21)	0.00-2.00	1.22 (1.45)	0.00-2.00
perceived Noise (0-6)	1.65 (1.48)	1.00-2.00	1.62 (1.44)	1.00-2.00	1.96 (1.58)	1.00-3.00
perceived Air pollution (0-6)	1.83 (1.55)	1.00-3.00	1.82 (1.52)	1.00-3.00	2.17 (1.64)	1.00-3.00
Non-specific symptoms (0-32)	5.96 (5.24)	2.00-8.00	5.66 (5.00)	2.00-8.00	5.64 (4.93)	2.00-8.00
Sleep disturbances (0-100)	27.18 (14.71)	16.11-35.56	26.42 (14.28)	15.56-33.89	25.40 (14.26)	15.56-33.33
Respiratory symptoms (0-5)	0.48 (0.97)	0.00-1.00	0.44 (0.91)	0.00-1.00	0.40 (0.87)	0.00-0.00

SD=standard deviation, IQR=interquartile range, RF-EMF=radiofrequency electromagnetic fields, NO_2=nitrogen dioxide.

Correlations

Table 3 shows the Spearman correlations between modeled exposures (RF-EMF, noise, air pollutants), perceived exposures, and symptom-based health outcomes (non-specific symptoms, sleep disturbances and respiratory symptoms). Correlation clusters were identified among the three modeled exposures (r_{Sp} 0.18-0.41), between modeled and corresponding perceived exposures (r_{Sp} RF-EMF= 0.11, noise=0.40, NO_2= 0.34), among the three perceived exposures (r_{Sp} 0.42-0.76), and the three health outcomes (r_{Sp} 0.27-0.50).

Table 3. Spearman correlation coefficients for modeled exposures, perceived exposure and symptom-based health outcomes in the AMIGO baseline cohort (n=14829, T0 = 2011/2012).

		Modeled exposure			Perceived exposure			Symptom-based health outcomes		
		RF-EMF (mW/m^2)	Noise (dB)	NO_2 ($\mu g/m^3$)	Base station	Noise	Air pollution	Non-specific symptoms	Sleep disturbances	Respiratory symptoms
Modeled exposure	RF-EMF (mW/m^2)									
	Noise (dB)	0.18								
	NO_2 ($\mu g/m^3$)	0.39	0.41							
Perceived exposure	Base station	0.11	0.11	0.15						
	Noise	0.14	0.40	0.28	0.42					
	Air pollution	0.15	0.35	0.34	0.46	0.76				
Symptom-based health outcomes	Non-specific symptoms	0.03	0.03	0.05	0.10	0.12	0.10			
	Sleep disturbances	0.03	0.03	0.06	0.13	0.14	0.12	0.50		
	Respiratory symptoms	0.03	0.03	0.05	0.06	0.07	0.08	0.37	0.27	

All correlations are significant with p-values <0.005.
RF-EMF=radiofrequency electromagnetic fields, NO_2=nitrogen dioxide.
Darker colors indicate higher correlations.

Effects of modeled and perceived exposure on symptom-based health outcomes

Table 4 summarizes the results of mixed model analyses of all perceived and modeled exposures and the different health outcomes. Modeled RF-EMF exposure was not significantly associated with respiratory symptoms and sleep disturbances, but was associated with lower non-specific symptom score in the single-predictor model. Perceived RF-EMF exposure was significantly associated with worse health outcomes in all single- and two-predictor analyses.

Modeled noise exposure was significantly associated with worse scores on health outcomes in the single-predictor models. Modeled noise exposure was associated with less sleep disturbance in the two-predictor model. Perceived noise exposure was significantly associated with worse health outcomes in all single- and two-predictor analyses.

Modeled NO_2 was significantly associated with worse scores on each health outcome in the single predictor models and in the two-predictor models, although effects of NO_2 diminished when perceived exposure was included in the two-predictor model. Perceived exposure to air pollution from road traffic was significantly associated with worse health outcomes in all single- and two-predictor analyses. Results for NO_x, $PM_{2.5}$, and PM_{10} were similar (Appendix Table A1), although the majority of the associations for these modeled air pollutants were not significant in the two-predictor models.

Table 5 summarizes the results of the fixed effect analyses in which temporal changes on an individual basis (between T0 and T1) in perceived exposure were related to changes in symptom reporting for the follow-up sample (n=7,905). For all environmental exposures, changes in perceived exposure were significantly associated with corresponding change in non-specific symptoms. Change in perceived RF-EMF exposure from base stations and noise exposure was significantly associated with a corresponding change in sleep disturbance. Change in perceived air pollution from road traffic was significantly associated with a corresponding change in respiratory symptoms.

Table 4. Mixed model analyses of Modeled and Perceived Exposure to RF-EMF from Mobile Phone Base Stations, Traffic Noise and Road Traffic Air Pollution on **Non-specific symptoms**, **Sleep disturbances**, and **Respiratory symptoms** for AMIGO respondents (n=14,829, T0 = 2011/2012 and n=7,905, T1=2015).

		Non-specific symptoms (0-32)		Sleep disturbances (0-100)		Respiratory symptoms (0-5)	
		β (95%CI)[a]	p	β (95%CI)[a]	p	β (95%CI)[a]	p
RF-EMF							
1	modeled (0-1)	-0.23 (-0.43,-0.03)	0.026[b]	-0.58 (-1.15,0.00)	0.051	-0.03 (-0.07,0.01)	0.096
2	perceived (0-6)	**0.37 (0.32,0.40)**	**0.000**	**0.81 (0.68,0.94)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (0-1)	-0.13 (-0.33,0.07)	0.201	-0.36 (-0.94,0.22)	0.222	-0.02 (-0.06,0.02)	0.305
	perceived (0-6)	**0.37 (0.32,0.41)**	**0.000**	**0.80(0.67,0.93)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
Noise							
1	modeled (dB)	**0.02 (0.01,0.03)**	**0.001**	**0.05 (0.01,0.09)**	**0.008**	**0.00 (0.00,0.01)**	**0.002**
2	perceived (0-6)	**0.30 (0.26,0.35)**	**0.000**	**0.83 (0.72,0.95)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (dB)	-0.01 (-0.03,0.00)	0.067	-0.04 (-0.08,-0.00)	0.028[b]	-0.00 (-0.00,0.00)	0.655
	perceived (0-6)	**0.32 (0.28,0.36)**	**0.000**	**0.88 (0.76,1.01)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
NO_2							
1	modeled (µg/m^3)	**0.05 (0.04,0.06)**	**0.000**	**0.15 (0.11,0.19)**	**0.000**	**0.01 (0.01,0.01)**	**0.000**
2	perceived (0-6)	**0.27 (0.23,0.31)**	**0.000**	**0.67 (0.56,0.78)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (µg/m^3)	**0.02 (0.01,0.04)**	**0.001**	**0.10 (0.05,0.14)**	**0.000**	**0.00 (0.00,0.01)**	**0.000**
	perceived (0-6)	**0.25 (0.21,0.29)**	**0.000**	**0.59 (0.48,0.71)**	**0.000**	**0.04 (0.03,0.06)**	**0.000**

1. These are the single predictor models for modeled exposure. 2. These are the single predictor models for perceived exposure. 3. These are the two-predictor models, i.e. including both modeled and perceived exposure.
[a]Adjusted for baseline values of sex, age, education, smoking, socioeconomic position, and year (baseline/follow-up).
[b]Beneficial effects with p-value below 0.05
Adverse effects are printed in bold if the p-value is lower than 0.05. RF-EMF=radiofrequency electromagnetic fields, PM= particulate matter, NO_2=nitrogen dioxide.

Effects of a change of environment

In total 1,224 (8.25%) participants moved to a different home between baseline in 2011/2012 (T0) and the follow-up questionnaire in 2015 (T1); of these, 592 participants filled in both questionnaires. This change of environment resulted in changed modeled and perceived exposures. Moved participants were categorized into three percentile based categories of change in absolute modeled exposure (decrease: 0-20 percentile, no change: 20-80, increase: 80-100). The cut-off points of the categories

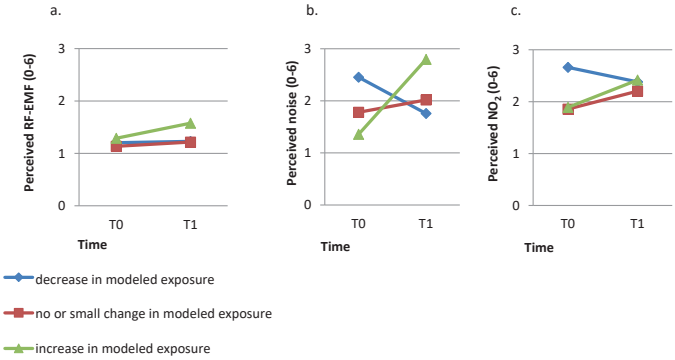
for the absolute change in modeled exposure are presented in Appendix Table A2. Figure 1 presents the course of mean perceived exposures over time for participant groups with a decrease, no or small change, and increase in modeled exposure after moving. For participants with an increase in modeled exposure, the corresponding average perceived exposure increased as well in the same time period, in particular for NO_2 and noise. For participants with a decrease in modeled exposure, the corresponding average perceived exposure decreased as well over time for noise and NO_2, but not for RF-EMF. Appendix Table A3 shows that intra-individual variation in perceived exposures and modeled exposures were not significantly associated with any intra-individual variation in symptom-based health outcomes, except for perceived RF-EMF, which was significantly associated with intra-individual variation in non-specific symptoms and sleep disturbance.

Table 5. Fixed effect analyses for effects of intra-individual changes in Perceived Exposure to Mobile Phone Base Stations, Noise and Air Pollution on intra-individual changes in **Non-specific symptoms** , **Sleep disturbances**, and **Respiratory symptoms** for AMIGO respondents (n=7,905) who participated at T0 (2011/2012) and T1 (2015).

	Non-specific symptoms (0-32)		**Sleep disturbances (0-100)**		**Respiratory symptoms (0-5)**	
	B (95%CI)	**p**	**B (95%CI)**	**p**	**B (95%CI)**	**p**
Perceived Base station (0-6)	**0.16 (0.10, 0.22)**	**0.000**	**0.19 (0.01, 0.36)**	**0.042**	0.01 (-0.00,0.02)	0.161
Perceived Noise (0-6)	**0.07 (0.01, 0.13)**	**0.021**	**0.21 (0.03, 0.39)**	**0.019**	0.01 (-0.01,0.02)	0.311
Perceived Air pollution (0-6)	0.04 (0.02, 0.10)	0.184	0.08 (-0.09, 0.25)	0.344	**0.01 (0.00,0.03)**	**0.049**

Adverse effects are printed in bold if the p-value is lower than 0.05.

Figure 1. Course of mean perceived exposures (a= RF-EMF, b =Noise, c= NO_2) over time (T0 =2011/2012, T1 = 2015) for AMIGO respondents who moved house between T0 and T1 (n=592) for percentile based categories (0-20, 20-80, 80-100) of absolute change in the corresponding modeled exposure.



* For each exposure, moved participants were divided in three percentile based categories (decrease: 0-20, no or small change in modeled exposure: 20-80, increase: 100) of the absolute change in modeled exposure between baseline and follow-up (see supplement table S3).
RF-EMF=radiofrequency electromagnetic fields and NO_2=nitrogen dioxide.

DISCUSSION

In this prospective cohort study, we examined the interplay between three modeled and perceived environmental exposures (RF-EMF from mobile phone base stations, noise and air pollutants from road traffic) and three symptom-based health outcomes (non-specific symptoms, sleep disturbances, and respiratory symptoms).

Correlation clusters

It seems that beliefs of participants about their exposure level to noise and air pollutants corresponded to some extent with their modeled exposure level, whereas this was not apparent for RF-EMF. In line with previous work[5], we found low correlations between modeled and perceived exposure to RF-EMF from mobile phone base stations. The low levels of knowledge regarding RF-EMF in the general population[43] likely plays a role. For example, the extent to which the exposure can be detected by the senses, and the visibility of nearby exposure sources. As expected, we found higher correlations between modeled and perceived exposure for noise exposure from traffic[1]. For air pollution from road traffic, correlations between modeled and perceived exposure were only slightly lower than for noise exposure. Perhaps familiarity with the link between road traffic and exposures, the visibility, noise, or smell of exhaust of nearby roads gave participant an indication of exposure near the home. We found moderate correlation clusters among the three modeled exposures, and substantial correlation clusters among the three perceived exposures and three different symptom-based health outcomes. As expected, we found correlations among modeled exposures, likely due to the clustering of exposures in urbanized areas. Correlation clusters among perceived exposures could be explained by a general environmental health worry factor[44], as well as the clustering of actual exposures. Correlations among health outcomes may be partly explained by underlying factor, representing a general tendency to report symptoms[45]. The presence of substantial correlation clusters among modeled exposures, perceived exposures, and health outcomes, implicates that disentanglement of different exposures and their individual health effects may prove difficult in epidemiological research.

Effects of modeled and perceived exposure on symptom-based health outcomes

Modeled RF-EMF was not associated with higher symptom scores, which is in line with earlier conducted studies[5,46,47]. For modeled noise exposure, prior studies on self-reported health[12–14] indicated that noise is mainly associated with increased sleep disturbances, and air pollutants mainly with respiratory symptoms[17–19]. The results of single predictor models in this study confirm the presence of significant adverse effects of noise and air pollutants on symptom scores. Contrary to our expectations, these

health effects extended across all assessed health outcomes, even those not previously reported in literature. However, the results were notably different in the two-predictor models, that included both modeled and perceived exposures. Significant adverse effects of modeled exposures on health outcomes generally disappeared (noise) or severely diminished (NO_2), when perceived exposure was included in the model. In two analyses (Table 4: effect of RF-EMF on non-specific symptoms, effect of noise on sleep disturbance) we found unexpected beneficial effects of modeled exposures, but these effects were small and possibly coincidental findings. The results do not necessarily imply the absence of causal effects of actual exposure on symptom scores, but do highlight the importance of exposure perceptions and a need to clarify the underlying causal mechanisms. The associations with symptom scores indicate a greater maximum impact of perceived than modeled exposure on symptom scores, for both single- and two-predictor models (as is shown in Appendix Table A4). These findings indicate that perceptions of exposures can play an important role when studying environmental determinants of symptom-based health outcomes.

A sizable minority of the participants reported high scores on perceived exposure levels. High scores on perceived exposures are likely to be in part the result of features of the environment that also drive modeled exposure levels (such as the proximity of nearby roads). In addition, worries about potential health effects of the specific exposure, and worries about environmental risks in general[25], can influence perceived exposure scores. A part of the cohort participants moved to a new home (n=592), and therefore changed their residential environment which often affected their modeled exposure levels. For this group, we found that a substantial increase or decrease in modeled exposure with respect to noise and air pollution (NO_2) was coupled with a simultaneous increase, respectively decrease in the corresponding perceived exposure (Figure 1). This longitudinal evidence strengthens the conclusion that participants are to some extent aware of, and able to estimate, the level of these two environmental exposures in their residential environment. The observed relation with change in perception was less distinct for RF-EMF from mobile phone base stations. Here, risk perception and health concerns appear to influence perceived exposure to a greater extent than exposure cues such as visibility of nearby base stations. In the group of follow-up participants (n=7,905), change in perceived exposures was significantly positively associated with change in most symptom-based health outcomes in the fixed effect analyses. This finding was not replicated in the smaller group of moved participants (n=592), except for positive effects of change in perceived RF-EMF on change in non-specific symptoms and sleep disturbance. However, due to the small number of movers, the power to detect such associations was limited in this subgroup. A change in perceived exposure in a new residential environment can be important given the associations between higher exposure perception and increased symptom scores, which were in line with earlier studies[3,4,48,49].

The implications of these findings in combination with the role of modeled exposures depend on the underlying causal mechanisms. A causal link from the exposure source both to modeled exposure (as a proxy of the true exposure level) and to perceived exposure is plausible[50], based on observability of exposure sources, and supported by the results of this study. For exposures that can be sensorially observed (f.i. noise) sensitivity and annoyance can play role as mediator[4] in the association between perceived exposure and symptom scores. In addition, there is sufficient evidence for the existence of nocebo effects[20,22,51], to support a causal link between perceived exposure and reported symptoms through negative health expectations when participants think they are exposed. If such nocebo effects occur in this population, mediation effects of exposure on symptom scores through perceived exposure would be likely. Such mediation mechanisms can have an impact on epidemiological studies examining environmental determinants of symptom-based health outcomes. When perceived exposure is not taken into account, indirect health effects through perceived exposure may be incorrectly ascribed to modeled exposures. However, the importance of such mediation mechanisms could be overestimated. Nocebo mechanisms have been mainly studied in laboratory and field-experiment studies, but the extent to which they are important for associations between perceived exposures and reported symptoms in the general population is unknown. Mechanisms of reversed causation may also play a role. For example, participants with health problems with an unknown cause may become more aware of exposures in their environment, and incorrectly start attributing these to environmental sources[10,52,53]. They experience and report their perceived exposure levels differently than healthy participants, which often is described as recall bias in epidemiological research and can be a problem in cross-sectional research and case-control studies. In this longitudinal study, with the use of a qualitative measure of perceived exposure, that is intended to capture the subjective experience of self-reported exposure, it perhaps is better described as a process of reversed causation. Depending on characteristics of the individual, but also features of the environment and changes in exposure, different processes underlying causal mechanisms of the link between exposure perceptions and symptom experiences could be important. Clarifying the underlying mechanisms is of great interest and importance for both epidemiological and psychosocial research disciplines, because of the implications for the interpretation of the relationships between the environment, perception, and symptom experiences. In addition, the need for effective public health intervention measures and policy implications varies depending on the importance of different mechanisms. Intervention measures targeted at reducing negative health expectations will only be effective in reducing symptom scores if the nocebo mechanism is the main explanation for the associations between exposure perception and reported symptoms.

Strengths and Limitations

The study had a large study population for studying the symptom-based health outcomes of interest. In addition, there were observations at two points in time, allowing for longitudinal analyses for a subset of participants. Thirdly, all studied environmental exposures were modeled using validated geospatial models that have been used in previous epidemiological research[37,39,40,54]. These models do not require extensive manual data-collection, allowing for research in large country-wide cohort studies. A limitation of the current study was that we only had modeled estimates for noise and air pollutants for one point in time (i.e. baseline), because input data for the geospatial models was not available for different years. Although estimates for noise and air pollutants would have improved slightly with new input data, large changes in exposure are not expected in this relatively short time frame. Eeftens et. al. (2011) showed that NO_2 decreased only slightly between 1997 and 2007 and correlations were high[55]. Another limitation concerns RF-EMF, where we modeled exposure from mobile phone base stations while the question about perceived exposure also included radio and tv base stations, because we expected people to not be familiar with differences between mobile phone and radio/tv base stations. However, given that mobile phone base stations are by far more present in residential areas than base stations for radio and tv, we expect these to dominate perceived base station levels.

Conclusion

Our study covered three environmental exposures, both modeled and perceived, and three symptom-based health outcomes. Correlations between modeled and perceived exposures appeared to be influenced by the observability of the exposure sources. Perceived exposures were consistently associated with increased symptom scores. In general, modeled exposures (except RF-EMF) were associated with increased symptom scores, but these associations disappeared or strongly diminished when perceived exposure was also added as a predictor. Under the reasonable assumption that perceived exposure is not a better proxy of the actual exposure than modeled exposure, these results would indicate that perceived exposure captures an additional element of the exposure that is not captured by the modeled exposure. When environmental determinants of symptoms are studied without acknowledging the potential role of these exposure perceptions, there is a risk of bias in the health effects attributed to modeled exposures. However, the etiological role of exposure perceptions in relation to symptom reporting requires further research. By combining insights from epidemiological and psychosocial research we have highlighted a range of complex issues that previously received little attention, but which can have important implications for interpretation of associations of interest, public health policy and intervention strategies.

Acknowledgments

This work was supported by The Netherlands Organisation for Health Research (ZonMw) within the programme Electromagnetic Fields and Health Research under grant number 85200001. The ESCAPE model used for air pollutant modelling received funding from the European Community's Seventh Framework Program (FP7/2007-2011) under grant agreement number 211250. The authors declare they have no actual or potential competing financial interests. We would like to thank Inka Pieterson for data management and Johan Beekhuizen for the original RF-modelling. We would also like to thank all the participants for taking the time and effort to participate in our study.

SUPPLEMANTARY TABLES CHAPTER 6

Table A1. Mixed model analyses of Modeled and Perceived Exposure to Air pollutants (NO_2, NO_X, $PM_{2.5}$, PM_{10}) on **Non-specific symptoms**, **Sleep disturbances**, and **Respiratory symptoms** for AMIGO respondents (n=14829, T0 = 2011/2012 and n=7905, T1=2015).

		Non-specific symptoms (0-32)		Sleep disturbances (0-100)		Respiratory symptoms (0-5)	
		β (95%CI)[a]	p	β (95%CI)[a]	p	β (95%CI)[a]	p
NO_2							
1	modeled (μg/m³)	**0.05 (0.04,0.06)**	**0.000**	**0.15 (0.11,0.19)**	**0.000**	**0.01 (0.01,0.01)**	**0.000**
2	perceived (0-6)	**0.27 (0.23,0.31)**	**0.000**	**0.67 (0.56,0.78)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (μg/m³)	**0.02 (0.01,0.04)**	**0.001**	**0.10 (0.05,0.14)**	**0.000**	**0.00 (0.00,0.01)**	**0.000**
	perceived (0-6)	**0.25 (0.21,0.29)**	**0.000**	**0.59 (0.48,0.71)**	**0.000**	**0.04 (0.03,0.06)**	**0.000**
NO_x							
1	modeled (μg/m³)	**0.02 (0.01,0.03)**	**0.000**	**0.05 (0.03,0.07)**	**0.000**	**0.00 (0.00,0.01)**	**0.000**
2	perceived (0-6)	**0.27 (0.23,0.31)**	**0.000**	**0.67 (0.56,0.78)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (μg/m³)	0.01 (-0.00,0.01)	0.153	0.01 (-0.01,0.04)	0.236	0.00 (-0.00,0.00)	0.068
	perceived (0-6)	**0.26 (0.22,0.30)**	**0.000**	**0.65 (0.53,0.77)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
$PM_{2.5}$							
1	modeled (μg/m³)	**0.20 (0.09,0.32)**	**0.001**	**0.59 (0.26,0.92)**	**0.000**	**0.02 (0.00,0.04)**	**0.041**
2	perceived (0-6)	**0.27 (0.23,0.31)**	**0.000**	**0.67 (0.56,0.78)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (μg/m³)	0.08(-0.04,0.19)	0.197	0.28 (-0.05,0.61)	0.096	0.00 (-0.02,0.02)	0.870
	perceived (0-6)	**0.27 (0.23,0.31)**	**0.000**	**0.66 (0.54,0.77)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
PM_{10}							
1	modeled (μg/m³)	**0.24 (0.16,0.32)**	**0.000**	**0.63 (0.40,0.86)**	**0.000**	**0.04 (0.02,0.05)**	**0.000**
2	perceived (0-6)	**0.27 (0.23,0.31)**	**0.000**	**0.67 (0.56,0.78)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**
3	modeled (μg/m³)	0.08 (-0.00,0.17)	0.056	**0.25 (0.01,0.49)**	**0.044**	0.01 (-0.00,0.03)	0.117
	perceived (0-6)	**0.26 (0.22,0.30)**	**0.000**	**0.64 (0.52,0.75)**	**0.000**	**0.04 (0.03,0.05)**	**0.000**

1. These are the single predictor models for modeled exposure. 2. These are the single predictor models for perceived exposure. 3. These are the two-predictor models, i.e. including both modeled and perceived exposure.
[a]Adjusted for baseline values of sex, age, education, smoking, socioeconomic position, and year.
Adverse effects are printed in bold if the p-value is lower than 0.05.
NO_2=nitrogen dioxide, NO_x = nitrogen oxide and PM=particulate matter.

Table A2. Mean, minimum and maximum absolute change in modeled exposures for the different percentile based categories (0-20, 20-80, 80-100) for AMIGO respondents (n=592) who moved to a different home address between T0 (2011/2012) and T1 (2015).

Exposure	Mean absolute change for each category	Minimum absolute change for each category	Maximum absolute change for each category
RF-EMF (mW/m^2)			
Decrease (percentile 0-20)	-0.078	-0.870	-0.010
No or small change (percentile 20-80)	0.003	-0.010	0.0280
Increase (percentile 80-100)	0.157	0.0290	1.417
Noise (dB)			
Decrease (percentile 0-20)	-10.308	-24.780	-6.100
No or small change (percentile 20-80)	-0.047	-6.100	5.400
Increase (percentile 80-100)	9.686	5.400	20.500
NO_2 (μg/m^3)			
Decrease (percentile 0-20)	-7.859	-27.382	-4.081
No or small change (percentile 20-80)	-0.371	-4.050	2.756
Increase (percentile 80-100)	6.872	2.770	31.779

RF-EMF=radiofrequency electromagnetic fields and NO_2=nitrogen dioxide

Table A3. Fixed effect analyses of effects of Modeled and Perceived Exposure to Mobile Phone Base Stations, Noise and Air Pollution on **Non-specific symptoms**, **Sleep disturbances**, and **Respiratory symptoms** for AMIGO respondents (n=592) that moved between baseline (2011/2012) and follow-up (2015) questionnaire.

		Non-specific symptoms (0-32)		Sleep disturbances (0-100)		Respiratory symptoms (0-5)	
		β (95%CI)	p	β (95%CI)	p	β (95%CI)	p
RF-EMF	Modeled (0-1)	0.42 (-0.26,1.11)	0.226	-0.03 (-2.05,1.99)	0.977	0.09 (-0.06,0.25)	0.237
	Perceived (0-6)	**0.30 (0.08,0.52)**	**0.008**	**0.71 (0.05,1.36)**	**0.034**	-0.00 (-0.05,0.05)	0.893
Noise	Modeled (dB)	-0.04 (-0.09,0.01)	0.100	-0.13 (-0.28,0.01)	0.074	-0.00 (-0.01,0.01)	0.526
	Perceived (0-6)	-0.03 (-0.22,0.16)	0.735	0.03 (-0.52,0.59)	0.902	-0.01 (-0.05,0.04)	0.777
NO_2	Modeled (μg/m^3)	-0.07 (-0.13,-0.01)	0.029[a]	-0.16 (-0.34,0.02)	0.079	-0.00 (-0.02,0.01)	0.499
	Perceived (0-6)	-0.08 (-0.27,0.11)	0.394	0.24 (-0.31,0.79)	0.394	-0.01 (-0.04,0.04)	0.926
NO_x	Modeled (μg/m^3)	-0.01 (-0.04,0.02)	0.548	-0.04 (-0.13,0.04)	0.306	0.00 (-0.00,0.01)	0.517
	Perceived (0-6)	-0.11 (-0.30,0.08)	0.262	0.21 (-0.35,0.76)	0.466	-0.01 (-0.05,0.03)	0.739
$PM_{2.5}$	Modeled (μg/m^3)	0.19 (-0.43,0.82)	0.541	-0.14 (-1.98,1.70)	0.880	-0.07 (-0.21,0.07)	0.311
	Perceived (0-6)	-0.13 (-0.32,0.06)	0.176	0.16 (-0.40,0.72)	0.571	-0.00 (-0.04,0.04)	0.995
PM_{10}	Modeled (μg/m^3)	-0.10 (-0.40,0.20)	0.519	-0.61 (-1.49,0.28)	0.180	-0.00 (-0.07,0.07)	0.992
	Perceived (0-6)	-0.10 (-0.29,0.09)	0.282	0.24 (-0.32,0.80)	0.402	-0.00 (-0.05,0.04)	0.835

RF-EMF=radiofrequency electromagnetic fields, NO_2=nitrogen dioxide, NO_x = nitrogen oxide and PM= particulate matter. Adverse effects are printed in bold if the p-value is lower than 0.05.
[a] beneficial effects with p-value below 0.05

Table A4. Comparison of effect sizes based on the Mixed effect model analyses of Modeled and Perceived Exposure to RF-EMF from Mobile Phone Base Stations, Traffic Noise and Road Traffic Air Pollution on **Non-specific symptoms**, **Sleep disturbances**, and **Respiratory symptoms** for AMIGO respondents (n=14829, T0 = 2011/2012 and n=7905, T1=2015).

		Non-specific symptoms (0-32)	**Sleep disturbances (0-100)**	**Respiratory symptoms (0-5)**
Predictors		maximum modeled effect on symptoms (maximum effect as percentage of the baseline mean)		
RF-EMF				
1	modeled (0-1)	-0.23 (-3.9%)	-0.58 (-2.1%)	-0.033 (-7.1%)
2	perceived (0-6)	2.21 (37.2%)	4.84 (17.8%)	0.042 (8.9%)
3	modeled (0-1)	-0.13 (-2.2%)	-0.36 (-1.3%)	-0.021 (-4.4%)
	perceived (0-6)	2.20 (37.0%)	4.81 (17.7%)	0.042 (8.9%)
Noise				
1	modeled (dB)	0.49 (8.2%)	1.16 (4.3%)	0.004 (0.8%)
2	perceived (0-6)	1.83 (30.8%)	5.00 (18.4%)	0.040 (8.6%)
3	modeled (dB)	-0.30 (-5.0%)	-1.01 (-3.7%)	-0.001 (-0.1%)
	perceived (0-6)	1.91 (32.2%)	5.29 (19.5%)	0.041 (8.7%)
NO_2				
1	modeled (µg/m^3)	1.09 (18.4%)	3.39 (12.5%)	0.009 (1.8%)
2	perceived (0-6)	1.62 (27.3%)	4.03 (14.8%)	0.043 (9.0%)
3	modeled (µg/m^3)	0.55 (9.2%)	2.10 (7.7%)	0.005 (1.0%)
	perceived (0-6)	1.50 (25.3%)	3.56 (13.1%)	0.038 (8.1%)
No_x				
1	modeled (µg/m^3)	0.82 (13.8%)	2.01 (7.4%)	0.004 (0.8%)
2	perceived (0-6)	1.62 (27.3%)	4.03 (14.8%)	0.043 (9.0%)
3	modeled (µg/m^3)	0.24 (4.0%)	0.55 (2.0%)	0.001 (0.3%)
	perceived (0-6)	1.57 (26.5%)	3.91 (14.4%)	0.040 (8.5%)
$PM_{2.5}$				
1	modeled (µg/m^3)	0.48 (8.1%)	1.40 (5.1%)	0.022 (4.6%)
2	perceived (0-6)	1.62 (27.3%)	4.03 (14.8%)	0.043 (9.0%)
3	modeled (µg/m^3)	0.18 (3.0%)	0.66 (2.4%)	0.002 (0.4%)
	perceived (0-6)	1.60 (26.9%)	3.93 (14.5%)	0.042 (9.0%)
PM_{10}				
1	modeled (µg/m^3)	0.83 (14.0%)	2.19 (8.1%)	0.037 (7.9%)
2	perceived (0-6)	1.62 (27.3%)	4.03 (14.8%)	0.043 (9.0%)
3	modeled (µg/m^3)	0.29 (4.8%)	0.85 (3.1%)	0.012 (2.7%)
	perceived (0-6)	1.55 (26.1%)	3.81 (14.0%)	0.040 (8.6%)

The maximum impact of the predictors on symptom scores was calculated by multiplying the range of the predictor values with the regression coefficients. The percentage in brackets represents the maximum impact of the predictors as a percentage of the mean symptom score at baseline. For continuous predictors the range was calculated as the 97.5th percentile minus the 2.5th percentile to exclude impact of extreme values.
RF-EMF=radiofrequency electromagnetic fields, NO_2=nitrogen dioxide, NO_x = nitrogen oxide and PM= particulate matter.

REFERENCES

1. Davies, H. W., Vlaanderen, J. J., Henderson, S. B. & Brauer, M. Correlation between co-exposures to noise and air pollution from traffic sources. *Occup. Environ. Med.* **66**, 347–350 (2009).
2. Allen, R. W. *et al.* The spatial relationship between traffic-generated air pollution and noise in 2 US cities. *Environ. Res.* **109**, 334–342 (2009).
3. Claeson, A.-S., Lidén, E., Nordin, M. & Nordin, S. The role of perceived pollution and health risk perception in annoyance and health symptoms: a population-based study of odorous air pollution. *Int. Arch. Occup. Environ. Health* **86**, 367–374 (2013).
4. Héritier, H. *et al.* The Association between Road Traffic Noise Exposure, Annoyance and Health-Related Quality of Life (HRQOL). *Int. J. Environ. Res. Public Health* **11**, 12652–12667 (2014).
5. Martens, A. L. *et al.* Modeled and Perceived Exposure to Radio-Frequency Electromagnetic Fields From Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort. *Am. J. Epidemiol.* 1–10 (2017) doi:10.1093/aje/kwx041.
6. Regel, S. J. *et al.* Pulsed radio-frequency electromagnetic fields: dose-dependent effects on sleep, the sleep EEG and cognitive performance. *J. Sleep Res.* **16**, 253–258 (2007).
7. Röösli, M., Frei, P., Mohler, E. & Hug, K. Systematic review on the health effects of exposure to radiofrequency electromagnetic fields from mobile phone base stations. *Bull. World Health Organ.* **88**, 887-896F (2010).
8. Baliatsas, C. *et al.* Non-specific physical symptoms and electromagnetic field exposure in the general population: Can we get more specific? A systematic review. *Environ. Int.* **41**, 15–28 (2012).
9. Hagström, M., Auranen, J. & Ekman, R. Electromagnetic hypersensitive Finns: Symptoms, perceived sources and treatments, a questionnaire study. *Pathophysiology* **20**, 117–122 (2013).
10. Dieudonné, M. Does electromagnetic hypersensitivity originate from nocebo responses? Indications from a qualitative study. *Bioelectromagnetics* **37**, 14–24 (2016).
11. Joseph, W. S. Night noise guidelines for Europe. *J. Am. Podiatr. Med. Assoc.* **100**, 1–162 (2009).
12. Pirrera, S., De Valck, E. & Cluydts, R. Nocturnal road traffic noise: A review on its assessment and consequences on sleep and health. *Environ. Int.* **36**, 492–498 (2010).
13. Öhrström, E. Sleep Disturbance , to Pilot Psycho-Social Survey Among of Road and Medical Exposed Traffic. *J. Sound Vib.* **133**, 117–128 (1989).
14. Muzet, A. Environmental noise, sleep and health. *Sleep Med. Rev.* **11**, 135–142 (2007).
15. Laszlo, H. E., McRobie, E. S., Stansfeld, S. A. & Hansell, A. L. Annoyance and other reaction measures to changes in noise exposure - A review. *Sci. Total Environ.* **435–436**, 551–562 (2012).
16. Kelly, F. J. Oxidative stress: its role in air pollution and adverse health effects. *Occup. Environ. Med.* **60**, 612–616 (2003).
17. Mar, T. F., Larson, T. V, Stier, R. A., Claiborn, C. & Koenig, J. Q. An Analysis of the Association Between Respiratory Symptoms in Subjects with Asthma and Daily Air Pollution in Spokane, Washington. *Inhal. Toxicol.* **16**, 809–815 (2004).
18. Modig, L. & Forsberg, B. Perceived annoyance and asthmatic symptoms in relation to vehicle exhaust levels outside home: a cross-sectional study. *Environ. Health* **6**, 29 (2007).
19. Patel, M. M. *et al.* Traffic-related particulate matter and acute respiratory symptoms among New York City area adolescents. *Environ. Health Perspect.* **118**, 1338–1343 (2010).
20. Szemerszky, R., Köteles, F., Lihi, R. & Bárdos, G. Polluted places or polluted minds? An experimental sham-exposure study on background psychological factors of symptom formation in 'Idiophatic Environmental Intolerance attributed to electromagnetic fields'. *Int. J. Hyg. Environ. Health* **213**, 387–94 (2010).

21. Crichton, F., Dodd, G., Schmid, G., Gamble, G. & Petrie, K. J. Can expectations produce symptoms from infrasound associated with wind turbines? *Health Psychol.* **33**, 360–4 (2014).
22. Witthöft, M. & Rubin, G. J. Are media warnings about the adverse health effects of modern life self-fulfilling? An experimental study on idiopathic environmental intolerance attributed to electromagnetic fields (IEI-EMF). *J. Psychosom. Res.* **74**, 206–212 (2013).
23. Hahn, R. A. The Nocebo Phenomenon: Concept, Evidence, and Implications for Public Health. *Prev. Med. (Baltim).* **26**, 607–611 (1997).
24. Köteles, F., Szemerszky, R., Freyler, A. & Bárdos, G. Somatosensory amplification as a possible source of subjective symptoms behind modern health worries. *Scand. J. Psychol.* **52**, 174–178 (2011).
25. Petrie, K. J. *et al.* Thoroughly modern worries: The relationship of worries about modernity to reported symptoms, health and medical care utilization. *J. Psychosom. Res.* **51**, 395–401 (2001).
26. Rief, W. *et al.* The relationship of modern health worries to depression, symptom reporting and quality of life in a general population survey. *J. Psychosom. Res.* **72**, 318–20 (2012).
27. Porsius, J. T. *et al.* Symptom reporting after the introduction of a new high-voltage power line: A prospective field study. *Environ. Res.* **138**, 112–117 (2015).
28. Spurgeon, A. Models of Unexplained Symptoms Associated with Occupational and Environmental Exposures. *Environ. Health Perspect.* **110**, 601–605 (2002).
29. Slottje, P., Yzermans, C. J., Korevaar, J. C., Hooiveld, M. & Vermeulen, R. C. H. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in The Netherlands. *BMJ Open* **4**, e005858 (2014).
30. Terluin, B. *et al.* The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* **6**, 34 (2006).
31. Hays, R. D., Martin, S. a, Sesti, A. M. & Spritzer, K. L. Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med.* **6**, 41–4 (2005).
32. Burney, P. G. J., Luczynska, C., Chinn, S. & Jarvis, D. The European Community Respiratory Health Survey. *Eur. Respir. J.* **7**, 954–960 (1994).
33. Sunyer, J. *et al.* Asthma score: Predictive ability and risk factors. *Allergy Eur. J. Allergy Clin. Immunol.* **62**, 142–148 (2007).
34. Kadaster Netherlands. Kadaster Netherlands. https://www.kadaster.nl/.
35. Bürgi, A., Theis, G., Siegenthaler, A. & Röösli, M. Exposure modeling of high-frequency electromagnetic fields. *J. Expo. Sci. Environ. Epidemiol.* **18**, 183–191 (2008).
36. Beekhuizen, J. *et al.* Modelling indoor electromagnetic fields (EMF) from mobile phone base stations for epidemiological studies. *Environ. Int.* **67**, 22–6 (2014).
37. Baliatsas, C., van Kamp, I., Swart, W., Hooiveld, M. & Yzermans, J. Noise sensitivity: Symptoms, health status, illness behavior and co-occurring environmental sensitivities. *Environ. Res.* **150**, 8–13 (2016).
38. Scheurs, E. M., Jabben, J. & Verheijen, E. N. G. *STAMINA-Model description STAMINA - Model description*. (2010).
39. Beelen, R. *et al.* Development of NO2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. *Atmos. Environ.* **72**, 10–23 (2013).
40. Eeftens, M. *et al.* Development of Land Use Regression Models for PM 2.5 , PM 2.5 Absorbance, PM 10 and PM coarse in 20 European Study Areas; Results of the ESCAPE Project. *Environ. Sci. Technol.* **46**, 11195–11205 (2012).
41. Statistics Netherlands. *Kerncijfers postcodegebieden* https://www.cbs.nl/ (2012).

42. Bell, A. *et al.* Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Polit. Sci. Res. Methods* **3**, 133–153 (2015).

43. Claassen, L., van Dongen, D. & Timmermans, D. R. M. Improving lay understanding of exposure to electromagnetic fields; the effect of information on perception of and responses to risk. *J. Risk Res.* **9877**, 1–17 (2015).

44. Petrie, K. J. *et al.* Worries about modernity predict symptom complaints after environmental pesticide spraying. *Psychosom. Med.* **67**, 778–782 (2005).

45. Porsius, J. T. *et al.* Somatic symptom reports in the general population: Application of a bi-factor model to the analysis of change. *J. Psychosom. Res.* **79**, 378–383 (2015).

46. Baliatsas, C. *et al.* Actual and perceived exposure to electromagnetic fields and non-specific physical symptoms : An epidemiological study based on self-reported data and electronic medical records. *Int. J. Hyg. Environ. Health* **218**, 331–344 (2015).

47. Frei, P. *et al.* Cohort study on the effects of everyday life radio frequency electromagnetic field exposure on non-specific symptoms and tinnitus. *Environ. Int.* **38**, 29–36 (2012).

48. Fyhri, A. & Klæboe, R. Road traffic noise, sensitivity, annoyance and self-reported health-A structural equation model exercise. *Environ. Int.* **35**, 91–97 (2009).

49. Frei, P., Mohler, E. & Röösli, M. Effect of nocturnal road traffic noise exposure and annoyance on objective and subjective sleep quality. *Int. J. Hyg. Environ. Health* **217**, 188–195 (2014).

50. Lima, M. L. On the influence of risk perception on mental health: living near an incinerator. *J. Environ. Psychol.* **24**, 71–84 (2004).

51. Porsius, J. T., Claassen, L., Woudenberg, F., Smid, T. & Timmermans, D. R. M. Nocebo responses to high-voltage power lines: Evidence from a prospective field study. *Sci. Total Environ.* **543**, 432–438 (2016).

52. Köteles, F. & Simor, P. Modern health worries, somatosensory amplification and subjective symptoms: A longitudinal study: A longitudinal study. *Int. J. Behav. Med.* **20**, 38–41 (2013).

53. de Graaff, M. B. (Bert). Should i be worried? Citizens' experiences and the risk politics of cell site deployment. (University of Amsterdam (UVA), 2016).

54. Martens, A. L. *et al.* Residential exposure to RF-EMF from mobile phone base stations: Model predictions versus personal and home measurements. *Sci. Total Environ.* **550**, 987–993 (2016).

55. Eeftens, M. *et al.* Stability of measured and modelled spatial contrasts in NO2 over time. *Occup. Environ. Med.* **68**, 765–770 (2011).

Chapter 7

General discussion

The overall aim of this thesis was to *expand the knowledge on possible associations between RF-EMF exposure from mobile phones and mobile phone base stations and health.* Firstly, we addressed several methodological issues which exist within epidemiological studies on health effects of RF-EMF exposure, such as setting up a large prospective cohort study on EMF exposure and health, improving exposure assessment of RF-EMF from mobile phone use by using both self-reported and operator-recorded data and improving outcome assessment of diseases such as Parkinson Disease. In addition, we looked at both actual, modeled, environmental exposure, and peoples perceptions about these exposures. In the last part of this thesis we explored associations between RF-EMF exposure from mobile phone and mobile phone base stations and health (several birth outcomes and self-reported symptoms).

In this thesis we used data collected from various data sources, such as data collected by mobile phone providers and data collected from various health care registries. Furthermore we combined data from different study populations in the prospective LIFEWORK cohort. These data provide ample new research opportunities to tackle some of the problems which exists within RF-EMF epidemiology today, such as the scarcity of longitudinal studies and the lack of a reliable exposure assessment method for RF-EMF form mobile phone base stations, that can be applied in large observational epidemiological studies. However, this data sources, such as mobile phone provider data and health care registry data also gives additional challenges regarding accuracy, validity and reliability of both exposure as well as health outcomes. In this final chapter, I will discuss the main findings of this thesis, which I will place in a broader context within the RF-EMF research field and opportunities and challenges of using multiple information sources.

Summary of the main findings

In the first two chapters we assessed and improved exposure assessment of mobile phone use in prospective cohort studies. In *Chapter 2* we described the LIFEWORK cohort in which we pooled data from three different prospective cohorts (Nightingale, EPIC-NL and AMIGO) to gain more statistical power and broader exposure distributions[1–4]. A description of the collected data, baseline characteristics, occupational and environmental exposure distributions are shown in this chapter. In total 88 466 participants were enrolled in three cohort studies between 2011 and 2012. To assess the reliability of self-reported data we compared information from participants who filled in the baseline questionnaire twice (n=237), in two different study populations. It showed that repeatability of the baseline questionnaire was high for stable factors such as educational level, but lower for items which exposure could change over time such as mobile phone use. *The main conclusion of this chapter was that we were able to pool data from three different cohort studies. Reliability was high for stable factors but lower for factors that change over time (mobile phone use). This prospective cohort offers a*

unique, large and rich resource for research on contemporary occupational and environmental health risks and will contribute to the large international COSMOS study on mobile phone use and health.

One of the main challenges in environmental epidemiology is accurate exposure assessment, as this is crucial for the power and validity of epidemiological studies. In *Chapter 3***,** we evaluated different statistical approaches (e.g. regression calibration, complete-case analyses, imputation) within the international prospective COSMOS study to combine operator-recorded and self-reported mobile phone use[5]. We showed that combining different sources (self-reported and operator-recorded) by regression calibration resulted in limited bias and outperformed other statistical methods such as complete-case and multiple imputation methods. In future exposure-health analyses within the COSMOS study the estimates of the regression calibration will be used to allow more accurate exposure-outcome relations for mobile phone use and health outcomes. *This study showed that regression calibration is an useful statistical method for combining multiple information sources as was shown for self-reported and operator-recorded mobile phone use. This method can be applied for other (environmental) exposures for which multiple data sources are available.*

Besides accurate exposure assessment, also accurate outcome assessment is important to obtain informative (unbiased and precise) risk estimates in epidemiological studies. The power of a study to detect associations between exposures and diseases will decrease when measurement error or under ascertainment of diseases occur[6]. Information on disease status is nowadays often collected from various data sources, such as registries and electronic medical records. In *Chapter 4* we combined multiple health information sources to develop a case ascertainment method for identifying Parkinson's Disease within two large Dutch cohort studies (AMIGO and EPIC-NL). Based on information from multiple data sources, such as health registries and questionnaires, we divided the participants into four categories of Parkinson Disease information, based on a developed likelihood score. We found that especially mortality records, ICPC codes from electronic medical records and self-reported PD diagnosis and PD medication had a high agreement (> 62%) with general practitioner information, while a screenings questionnaire (Tanner) and ATC codes from electronic medical records had lower agreement with general practitioner diagnosed PD (< 20%). Furthermore, we tested our likelihood score with known (strong) etiological risk factors for PD. For EPIC-NL the associations (smoking, sex and family history) for the highest category were in line with the literature. *Our study showed that performance of the different data sources varied in validity, therefore we recommend that for diseases for which no reliable registry is available, such as Parkinson Disease in the Netherlands, it remains important to include and combine data from multiple sources for case ascertainment.*

The previous three chapters assessed the quality of the exposure or health outcome data and helped improve assessment of both. In the second part of this thesis we investigated the associations between

RF-EMF exposure and health (several birth outcomes and self-reported symptoms). In *Chapter 5* we examined maternal self-reported mobile phone use in relation to several birth outcomes (pregnancy duration, birth weight, fetal growth) in four birth cohorts. We showed that pregnancy duration (lower gestational age/shorter pregnancy duration) and preterm birth might be associated with maternal mobile phone use during pregnancy, although results could also reflect stress during pregnancy or other residual confounding factors rather than the direct effect of RF-EMF emitted by mobile phones. We found no effects with birth weight or fetal growth. The results for pregnancy duration were rather stable among the four birth cohorts, although we found a reversed association in one of the cohorts. *Our study showed an indication for associations between pregnancy duration, preterm birth and mobile phone use, however it remains uncertain whether this is due to RF-EMF from mobile phone use or related to other factors, such as stress during pregnancy.*

In *Chapter 6* we reported associations between environmental exposures (RF-EMF from mobile phone base stations, road traffic noise and air pollutants) and several self-reported health outcomes (non-specific symptoms, sleep quality and respiratory health). In this study we included the 'actual' exposure, which was modeled on the home address, and people's perceptions about these environmental exposure. These perceived exposures showed stronger associations with the different health outcomes than the modeled exposures. Modeled air pollutants and road traffic noise were associated with higher symptom scores, but when accounting for their perceived exposure these associations disappeared or diminished. We found no effects of modeled RF-EMF exposure from mobile phone base stations on self-reported symptoms, while we found effects of perceived exposure of mobile phone base stations, which is in line with previous research[7,8]. *The main conclusion of this study is that no association was found for modeled RF-EMF exposure and associations for modeled road traffic noise and air pollutants disappeared or diminished when taken into account perceived exposure. Therefor both perceived and modeled exposure should be taken into account when studying (self-reported) symptoms, otherwise associations might be falsely attributed.*

RF-EMF epidemiology

The way people use their mobile communication devices has constantly changed since they were introduced in the early nineties. From only using a mobile phone for making calls at the introduction of the mobile phone to constantly browsing and streaming on a smartphone nowadays. In my thesis we looked at RF-EMF exposure from mobile phone use. At the time of the baseline COSMOS study (2006-2012) calling (self-reported and operator-recorded) was one of the main exposures sources of RF-EMF[9]. Nowadays, there are more sources emitting RF-EMF like tablets and Wi-Fi, however RF-EMF exposure from mobile phone use is still the main interest in RF-EMF research as exposure when calling is very localized at the head and high compared to other sources emitting RF-EMF[10].

Exposure assessment RF-EMF

Exposure assessment of RF-EMF is challenging as many factors will influence participants' total RF-EMF exposure. RF-EMF is emitted from multiple near- and far-field sources and depend on other aspects such as device position, timing, usage function, output powers and personal characteristics. First, I will discuss exposure assessment of RF-EMF from mobile phone use, followed by changes over time and combining near- and far-field exposure.

Exposure assessment of mobile phone use

In Chapter five maternal mobile phone use exposure was based on self-reported frequency of calls. In one of the four birth cohorts (MOCEH cohort), data was compared with more objective operator log data which showed moderately to high agreement[11]. In the older cohorts (DNBC, ABCD) the frequency of calls was reported retrospectively, 7 years after pregnancy, which could have influenced the quality of the reported use. The exposed participants within the older cohorts might be early adaptors of mobile phone having other characteristics than the population in general. To account for this, we performed sensitivity analysis on the data collected within the older (DNBC, ABCD) and newer cohorts (INMA, MOCEH). This showed that the results for post term birth were driven by the older cohorts. Within the cohorts DBCD, ABCD and INMA low exposed participants made 0-1 calls per day while high users made more than 4 calls per day. In the future ideally more detailed information or even information from other sources, such as operator-recorded data, would be preferred. Longitudinal analyses should be conducted to determine causality of the relationship, furthermore we should also take into account perception and stress during pregnancy.

Similar with earlier studies on exposure assessment of mobile phone use we observed in COSMOS (chapter 3) that high levels of self-reported mobile phone use were overestimated while low levels of mobile phone use were underestimated compared to operator-recorded data. Recall bias in self-reported mobile phone information is a well-known problem in the RF-EMF research field and can affect results by a decrease in study power and risk estimates, which will be biased towards the null[12]. Alternatives for self-reported data are operator-recorded data, software modified phones or smartphone applications. Software modified phones or applications, such as XMobisense, are measuring multiple aspects of mobile phone use including laterality and handsfree use. Data collected through these applications (such as XMobisense) or software modified phones gives us more detailed and accurate information than self-reported or operator-recorded data[13,14]. Some of these data are hard to capture in questionnaires, such as laterality or handsfree use, as agreement was shown to be moderate between self-reported information and information collected by smartphone applications[13,14]. Comparison of data from software modified phones differed with self-reported data and also showed poor agreement with self-reported data use[10]. Although information from software

modified phones or smartphone applications is more objective then self-reported information the applicability in larger study population is more difficult. This was shown by a small pilot study, which we performed within AMIGO, investigating the feasibility to include the XMobisense application in an ongoing cohort study. The response rates were rather low, maybe perhaps due to privacy concerns and low information density for the users themselves. Some participants experienced problems with installation of the application and after a couple of months most participants removed the application from their device again, so for prospective research active follow-up will be needed to make sure the application stays installed on people's device. In conclusion, collecting data by smartphone application or software modified phones is more feasible in smaller research settings, for example for validation purposes, rather than using it to collect data for a large amount of participants in prospective cohorts.

Another exposure assessment method to include in larger (prospective) studies is operator-recorded mobile phone data collected by mobile phone providers. While operator-recorded data in most studies is regarded as the golden standard, this more objective measure of mobile phone use also has several shortcomings. Data extraction and linkage to provider data is complex. For example not everyone is willing to give informed consent to access these data and share their mobile phone number. For a proportion of participants with informed consent it is not possible to link their mobile phone number to the actual data because issues occur regarding ownership of the mobile phone number such as not matching personal information (e.g. name on another (company) name) or prepaid mobile phone services. Furthermore, mobile phone providers collect data for billing purposes rather than research purposes, which have an influence on the data collected. For example in the billing system of some providers only outgoing calls are registered, because consumers only need to pay their outgoing calls. However for research, exposure to both incoming and outgoing calls are relevant. Operator-recorded data will not always be complete and this can also differ between providers or depend on personal characteristics. Furthermore mobile phone providers often cooperate voluntarily and therefore it is important to keep providers involved and willing to share their data.

Although we know that participants have problems with recalling their device use resulting in either under- or overestimation of use, collecting self-reported information is still an important method in large population studies. In chapter 3 we have shown that self-reported information can be used in combination with operator-recorded data in a smaller subset of participants by applying statistical methods such as regression calibration. With self-reported information it is relatively easy to collect data on a wide variety of aspects and sources such as historical and current mobile phone use, both frequency and duration of calls and the use of other personal devices such as tablets and wearables which are important when investigating peoples total RF-EMF exposure. If we can adjust for bias, for example by regression calibration, questionnaires are an useful method to collect data on especially near-field EMF sources.

For far-field sources, such as mobile phone base station, Wi-Fi and broadcasting antennas, approaches with questionnaires are less suitable. In chapter 6 we showed that Spearman correlation between modeled RF-EMF exposure from base station and perceived exposure showed only a very slight agreement (Rsp= 0.11). Other studies looking at the agreement between modeled and self-reported exposure from mobile phone base station also showed low agreements[7].

Useful tools to calculate exposure from RF-EMF exposure from base stations are geographical information systems, such as the NISMap propagation model used in this dissertation [15–17]. This model can estimate RF-EMF exposure levels at any location using input data on antenna location, building characteristics and transmission powers[18,19]. Validation studies showed a moderate agreement (Spearman correlation) between modeled RF-EMF and measured RF-EMF with personal monitors[20,21]. Feasibility in large prospective studies is no problem as (self-reported) home address are available in (prospective) studies on environmental exposures. However, exposure assessment of RF-EMF from mobile phone base stations by propagation models will become more challenging with the introduction of the 5G technology which will be hard to model due to the directed beams, and an increase of micro and femtocells implemented on a local scale.

Another method to measure far-field EMF sources, often used in smaller studies, is the use of personal exposure monitors[22,23]. Personal exposure monitors will be hard to implement in larger observational studies because they are expensive and the monitors will need frequent calibration.

Changes over time

Usage of mobile phones has enormously increased over time, however RF-EMF emitted from mobile phones has decreased over time with the adaption of third-generation (3G) devices, which have a much lower output power than second-generation (2G) devices[24]. This has continued with the implementation of the more recent technologies (4G) although differences with 3G are less in terms of users' exposure. With the recent introduction of the 5G technology RF-EMF exposure emitted by mobile phone will change again. The implementation of newer communication technologies is implemented gradually and mobile phones can switch to older network generations which will make it hard to exactly determine exposure levels and participants' life time exposure. Estimating exact exposure levels for each individual is not necessary for epidemiological research when looking at associations, but a valid relative ranking of exposed individuals without misclassification is essential[12].

Combining near- and far-field RF-EMF exposure

Lately more attention has been given to combining RF-EMF from both near-field (less than one wavelength away from antenna) and far-field (two or three wavelength away) sources. Exposure to a single source might be low, but an individual's total RF-EMF exposure can be substantial when

combining all sources. Van Wel and colleagues recently developed an Integrated Exposure Model[10] (IEM), including ten different near- and far-field sources emitting RF-EMF. This IEM showed that for outcomes near the brain the exposure from the mobile phone is still the main exposure source, while for the other organ-systems the contribution of other sources will contribute relatively more to the total RF-EMF exposure[10]. Until now most epidemiological studies investigating effects of RF-EMF exposure solely used (self-reported) mobile phone use, therefor it would be good to implement such an Integrated Exposure Model in ongoing cohort studies, such as COSMOS. It is essential that these models are regularly updated due to the fast technological developments, such as the current implementation of 5G in the Netherlands and elsewhere, and changes in the utilization of the devices.

Implications for future research

Straightforward exposure assessment approaches such as only assessing mobile phone use at one point in time will likely result in exposure misclassification for some organ-systems. Future studies should focus on implementing newly developed Integrative Exposure Models (IEM), such as developed by van Wel et al[10]. To include an IEM model it is important for cohort studies to include all relevant sources emitting RF-EMF in questionnaires. Currently, large epidemiological studies often lack one or more sources needed as input for an IEM model. For most sources such as tablets, wearables and Wi-Fi it still has to be determined how well participants can estimate their use. Validation studies are absent for most of these sources as research mainly focused on validation exposures from mobile phone use or mobile phone base stations[13,25–30]. Validation of input data for the IEM and propagation models is necessary to reduce measurement error and to gather information on uncertainties in the estimated exposure. Regression calibration exercises, as shown in chapter 3, can help to overcome problems when self-reported information has measurement error and operator-recorded or smartphone application data is not available for the entire study population (over time).

Given the rapidly changing technologies and utilization of mobile communications devices, research on RF-EMF will likely be outpaced by the adoption of these new technologies. For example, the 5G is technologically different from previous generations by using other (higher) frequency bands and directed beams at a local place which has a influence on the current exposure models used in epidemiological research. Therefor it remains important to keep the research up to date, so physical consequences of mobile communication devises can be captured.

Perceived exposure, behavioral aspects and stress

Although the evidence for a causal relation between RF-EMF exposure and health is still unclear, mobile phone and other sources emitting RF-EMF such as mobile phone base station could also indirectly affect health. We found that people can experience symptoms when they think they are

exposed to RF-EMF from mobile phone base stations regardless of the actual, modeled, exposure level as has been shown in chapter 6 of this thesis. The findings for mobile phone base stations could also be relevant for mobile phones or other devices emitting RF-EMF especially if people have concerns related to these technologies. Furthermore in chapter 5 we could not rule out confounding of the results found between maternal mobile phone use and pregnancy duration as we were unable to correct for stress during the pregnancy. Also more recent results from the large international COSMOS study have suggested that stress and addictive behavior can play an important role in symptoms such as headache and sleep quality[31,32]. The use of mobile communication devices before sleeping can have a negative effect on sleep-quality, reason for this could be the blue-light which the devices emit or the delay or interruption of sleep time by prolonged use of devices[33].

In chapter 6 we showed that perceived exposure can cause an increase in symptoms regardless of the modeled level of exposure. We found this for three different environmental exposures: road traffic noise, different air pollutant measures and RF-EMF from mobile phone base stations. The findings from chapter 6 confirms earlier research which showed that perception is sometimes even stronger associated with (self-reported) symptoms than actual (modeled) exposure levels[7,8,34–36]. Based on our findings our recommendation is to take perceived exposure into account as an independent variable in addition to the actual physical component of environmental exposures when researching symptoms otherwise effects can be falsely attributed to the modeled exposure. Various mechanism can play a role in these findings, including the nocebo process (expectation that negative health effect may occur can have an adverse impact on symptom reporting) or participants with many symptoms may monitor their environment more actively and recall their exposure differently than healthy participants. The results from chapter 6 showed that participants could estimate their exposure to road traffic noise and air pollution better than their exposure to RF-EMF from mobile phone base stations. The agreement between modeled and perceived exposure for mobile phone base station was rather poor with a Spearman correlation of 0.11, which indicate that people cannot estimate their exposure to RF-EMF from mobile phone base stations very well. A possible explanation that correlations between modeled and perceived exposures were higher for road traffic noise and air pollutants is that the common exposure source “traffic” is more visible and RF-EMF exposure is invisible and propagation patterns are more complex.

The different perceived exposures were highly correlated in our study (Spearman correlations between 0.42 and 0.76). This could be due to the actual clustering of environmental exposures in certain areas (for example in urban or rural regions) as also moderate to high correlation between modeled exposures were shown (Spearman correlations between 0.18-0.41), or because of more general worries about environmental exposures which work through the same psychosomatic mechanism. These worries might be driven by for example attention in social media, as currently is the case for

the 5G technology of mobile phone base stations. For modern health worries several association are shown with self-reported health outcomes[37].

Not only are individuals simultaneously exposed to a complex mixtures of multiple environmental and occupational exposures, our research also showed that next to the actual physical component of exposure also perception can influence people's health. The role of people's perception and the social environment of people should get more attention in future initiatives around the exposome and within RF-EMF epidemiology.

Implications for future research

Nowadays everybody is exposed to RF-EMF, so even a possible small risk, such as found in chapter 5 for birth outcomes, may, if causal, translate into public health problems. The exact biological mechanism for the results found in this thesis are still unclear and could be related to stress or addictive behavior. Furthermore we showed that perceived exposure to RF-EMF from mobile phone base stations is related to symptom reporting. More focus should be given on researching potential psychosocial aspects of devices emitting RF-EMF to disentangle if health effects are related to the electromagnetic fields or are related to other aspects of mobile communication devices, such as stress, worries, always being available and addictive behavior of devices. If the latter would be true, more emphasis should be given to risk communication strategies rather than exposure reduction strategies.

For the associations between perceived exposure and symptoms, there is still a lack of clarity how these psychosocial mechanisms work. Although experimental studies can help clarifying the role of different psychosocial mechanism there is also a need to study perceived exposures and symptom reporting in more longitudinal studies. Currently data is often only available in cross-sectional studies or at most for two points in time. With repeated measures changes over time in exposure perceptions and symptom reporting can help to understand the causality of the relation and help to answer the questions what comes first? Within AMIGO, we made a start with investigating this in a longitudinal way as has been shown in this thesis and by the work of Martens et al.[8,38,39]. This showed bidirectional temporal associations between health concerns and symptoms. Furthermore an increase in modeled exposure to RF-EMF from mobile phone use was associated with an increase in perceived exposure.

Multiple sources of information

Within different chapters (2-4) of this thesis we showed how data from different sources can be pooled and combined to improve research within RF-EMF epidemiology. This can be done prospectively such as conducted within COSMOS and LIFEWORK, or retrospectively as has been performed within GERoNiMO, an European project on Generalized EMF research using novel methods. Furthermore

we assessed the quality of different data sources, for example mobile phone use in chapter 3 and Parkinson Disease in chapter 4. Although the examples in this thesis were rather specific, findings can be applicable to other study populations, exposures or health outcomes. We combined data on an individual level (chapters 2-4) or through meta-analyses (chapter 5) with the main advantage that we had a larger sample size and a richer dataset available to study multiple health outcomes and exposures. Replication of findings in (environmental) epidemiology is important to minimize the chance of false or spurious findings. However, more power gained due to combining studies often goes together with less detail in exposure assessment. In this paragraph I will discuss the opportunities and challenges of combining multiple study populations and using multiple data sources for outcome and exposure assessment based on our experiences within this thesis.

For the LIFEWORK cohort (chapter 2) we used a harmonized questionnaire to characterize exposure and health in different cohorts (EPIC-NL, AMIGO, Nightingale), however the setup and aims of each cohort, including the study population, was different. The same was true for the international COSMOS study which data was analyzed in chapter 3. Within the COSMOS study the recruitment methods, study populations and timing of the baseline questionnaire differed between the countries, but the questionnaire on mobile phone use was largely harmonized beforehand. In chapter 5 we pooled data from four birth cohorts for which we performed meta-analyses on the association between mobile phone use during pregnancy and several birth outcomes. Comparison of the data within chapter 3 and 5 was difficult as exposure to RF-EMF and use of mobile phone has changed enormously within a short time frame.

When pooling data unfortunately the amount of detail gets less when keeping variables comparable among the different study populations and countries. Examples can be found in exposure variables but also in outcome and confounder variables. Variables could be non comparable between countries (e.g. educational level) or different answer categories were used in the conduct of the questionnaire (e.g. mobile phone use). As a result during the standardization process detailed information needs to be collapsed when combined with less detailed information from another study to keep results comparable. When conducting a meta-analyses, variables used for confounding could be optimized within each cohort and remain slightly different in the different study populations. For example in the COSMOS study the variable educational level was asked differently due to differences in educational system in the countries, therefore we were only able to use a categorical variable with two categories (low versus high), while in most countries more than 5 categories were available.

Standardization is not only an issue with self-reported data but also for other data sources. For example operator-recorded mobile phone use was collected differently in the COSMOS study; in Sweden, Finland and UK information was available per call while in Denmark and the Netherlands

monthly aggregated data was received from providers. Therefor we had to aggregate data for Sweden, Finland and UK. Although we tried to standardize protocols as much as possible at the start of the projects, it is inevitable that there are still variables that need to be harmonized after data collection. Furthermore, we noticed differences in the way the data is collected by the providers. For example in the operator-recorded mobile phone data we noticed that in some countries data on incoming calls between the same providers were not recorded while in other countries they were included. This has led to differences in the ratio between incoming and outgoing calls which we expected to be similar across the different providers and countries. Therefore we decided in chapter 3 to conducted the analyses on outgoing mobile phone calls as the primary exposure.

More data and performing analyses in multiple countries and study populations could lead to contradicting risk estimates. There could be underlying reasons for the differences found such as collecting data over various years, quality of the data or differences between study populations. Results should be interpreted with caution and replicated in other studies. For example in chapter 3, exposure distributions varied between the countries. Within the COSMOS study the baseline questionnaire in Denmark was collected already in 2007 while in the Netherlands and UK the first questionnaire was collected between 2011 and 2012. Within these 5 years the exposure patterns of RF-EMF changed enormously by the introduction of novel network generations (≥3G), introduction of smartphones and changes in use. This could have influenced the results and should be taken into account in future analyses, for example as has been done in the COSMOS paper about headache[32]. In this article a distinction was made between calls made on the 2G and 3G network. In chapter 5, we also performed additional sensitivity analyses, to account for network generations, by making a distinction between the more recent and older cohorts[40].

For PD identification we showed, in chapter 4, that agreements differed for PD identification based on ATC codes or ICPC codes from electronical medical records. This difference could be explained by the fact that medication for PD is often used as a diagnostic tool for Parkinson's Disease. As PD is regarded as a chronic disease the prescription remains in the EMR system leading to a false indication of PD when using ATC codes for PD diagnosis. EMR are designed for healthcare purpose rather than for scientific purposes and therefor the data should be used with care. If not assessing validity of sources these differences could be missed. Recently more research is making use from electronic healthcare records so it remains important to accurately process the data from electronic records to ensure the quality of the data[41].

Implications for future research

With the increase in sharing, harmonizing, pooling and standardizing data it remains important to fully understand the data we are using. As we nowadays combine data from various information sources we need to understand the quality of the data that is used, therefor it remains important to assess accuracy, validity and reliability of the sources. This is true for both outcome as exposure data, as shown in this thesis for Parkinson's Disease and mobile phone use. Validation studies remain of major importance. Investments should be made for standardizing diseases in national registries. With the shift towards more integrative exposure models or investigating participants' exposome it is essential that the input data of the different sources is of sufficient quality. Otherwise statistical issues will occur such as making wrong inferences because variables with the largest measurement error might be forced out of the model. Within the harmonization and standardization process it is important to avoid the lowest denominator, because if too much detail is lost results could become uninformative.

Concluding remarks

On the one hand the results presented in this thesis add to improvements in the assessment of mobile phone use and Parkinson's Disease for future epidemiological investigations. We showed (statistical) approaches to combine data from various sources, which could be applied on other exposure and outcomes as well. Nowadays the research community is making more use of 'big data' so these aspects will probably become even more important in the future. Our work highlights the importance of determining the validity of the different exposure and health measures when using multiple sources within environmental epidemiology. In the future the regression calibration results will be used in health-outcome analysis within the COSMOS study. Furthermore Parkinson's Disease will be investigated in relation to several environmental exposures, including road traffic noise and air pollutants, within AMIGO and EPIC-NL. The results in this thesis add to the existing body of evidence for a relationship between RF-EMF exposure and health. We found indications that maternal mobile phone use during pregnancy might be associated with preterm birth and shorter pregnancy duration. However interpretations of these findings is still difficult as we were unable to account for stress during pregnancy. Furthermore we showed the importance of people's perceptions in symptom reporting for not only RF-EMF exposure from mobile phone base stations but also for road traffic noise and air pollutants. This emphasizes the need for multidisciplinary research by including insights both from the epidemiological as the psychosocial perspective. This thesis provided an insight on how people's health can be affected by their living environment, both by the exposure itself as by people's perception, especially for RF-EMF.

REFERENCES

1. Reedijk, M. *et al.* Cohort profile : LIFEWORK , a prospective cohort study on occupational and environmental risk factors and health in the Netherlands. (2018) doi:10.1136/bmjopen-2017-018504.
2. Pijpe, A. *et al.* The Nightingale study: Rationale, study design and baseline characteristics of a prospective cohort study on shift work and breast cancer risk among nurses. *BMC Cancer* **14**, (2014).
3. Beulens, J. W. J. *et al.* Cohort profile: the EPIC-NL study. *Int. J. Epidemiol.* **39**, 1170–8 (2010).
4. Slottje, P., Yzermans, C. J., Korevaar, J. C., Hooiveld, M. & Vermeulen, R. C. H. The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in the Netherlands. *BMJ Open* **4**, e005858 (2014).
5. Schüz, J. *et al.* An international prospective cohort study of mobile phone users and health (Cosmos): design considerations and enrolment. *Cancer Epidemiol.* **35**, 37–43 (2011).
6. Rothman, K. J. & Health, P. Methodologic Frontiers Epidemiology in Environmental Introduction. **101**, 19–21 (2011).
7. Baliatsas, C. *et al.* Actual and perceived exposure to electromagnetic fields and non-specific physical symptoms: An epidemiological study based on self-reported data and electronic medical records. *Int. J. Hyg. Environ. Health* **218**, 331–344 (2015).
8. Martens, A. L. *et al.* Modeled and Perceived Exposure to Radiofrequency Electromagnetic Fields From Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort. *Am. J. Epidemiol.* **186**, 210–219 (2017).
9. Bolte, J. F. B. & Eikelboom, T. Personal radiofrequency electromagnetic field measurements in the Netherlands: Exposure level and variability for everyday activities, times of day and types of area. *Environ. Int.* **48**, 133–142 (2012).
10. van Wel, L. *et al.* Organ-specific integrative exposure assessment: Radiofrequency electromagnetic fields exposure and contribution of sources in the general population. (2020).
11. Choi, K. *et al.* MULTIPLE ASSESSMENT METHODS OF PRENATAL EXPOSURE TO RADIO FREQUENCY RADIATION FROM TELECOMMUNICATION IN THE MOTHERS AND CHILDREN ' S ENVIRONMENTAL HEALTH (MOCEH) STUDY. **29**, 959–972 (2016).
12. Roosli, M. *Epidemiology of electromagnetic fields* . (2014).
13. Goedhart, G., Kromhout, H., Wiart, J. & Vermeulen, R. Validating self-reported mobile phone use in adults using a newly developed smartphone application. *Occup. Environ. Med.* **72**, oemed-2015-102808 (2015).
14. Goedhart, G. *et al.* Recall of mobile phone usage and laterality in young people: The multinational Mobi-Expo study. *Environ. Res.* **165**, 150–157 (2018).
15. Bürgi, A. *et al.* A model for radiofrequency electromagnetic field predictions at outdoor and indoor locations in the context of epidemiological research. *Bioelectromagnetics* **31**, 226–236 (2010).
16. Bürgi, A., Theis, G., Siegenthaler, A. & Röösli, M. Exposure modeling of high-frequency electromagnetic fields. *J. Expo. Sci. Environ. Epidemiol.* **18**, 183–91 (2008).
17. Frei, P. & Röösli, M. Exposure to radiofrequency electromagnetic fields in our everyday environment. in *Epidemiology of Electromagnetic Fields* 125–140 (2014). doi:10.1201/b16968.
18. Beekhuizen, J., Vermeulen, R., Kromhout, H., Bürgi, A. & Huss, A. Geospatial modelling of electromagnetic fields from mobile phone base stations. *Sci. Total Environ.* **445–446**, 202–209 (2013).
19. Beekhuizen, J., Kromhout, H., Bürgi, A., Huss, A. & Vermeulen, R. What input data are needed to accurately model electromagnetic fields from mobile phone base stations? *J. Expo. Sci. Environ. Epidemiol.* **25**, 53–57 (2015).

20. Martens, A. L. *et al.* Residential exposure to RF-EMF from mobile phone base stations: Model predictions versus personal and home measurements. *Sci. Total Environ.* **550**, 987–993 (2016).

21. Martens, A. L. *et al.* Validity of at home model predictions as a proxy for personal exposure to radiofrequency electromagnetic fields from mobile phone base stations. *Environ. Res.* **142**, 221–226 (2015).

22. Aminzadeh, R. *et al.* The effect of antenna polarization and body morphology on the measurement uncertainty of a wearable multi-band distributed exposure meter. *Ann. des Telecommun. Telecommun.* **74**, 67–77 (2019).

23. Frei, P. *et al.* Classification of personal exposure to radio frequency electromagnetic fields (RF-EMF) for epidemiological research: Evaluation of different exposure assessment methods. *Environ. Int.* **36**, 714–720 (2010).

24. Lauer, O. *et al.* Combining near- and far-field exposure for an organ-specific and whole-body RF-EMF proxy for epidemiological research: A reference case. *Bioelectromagnetics* **34**, 366–374 (2013).

25. Vrijheid, M. *et al.* Recall bias in the assessment of exposure to mobile phones. *J. Expo. Sci. Environ. Epidemiol.* **19**, 369–381 (2009).

26. Heinävaara, S., Tokola, K., Kurttio, P. & Auvinen, A. Validation of exposure assessment and assessment of recruitment methods for a prospective cohort study of mobile phone users (COSMOS) in Finland: a pilot study. *Environ. Health* **10**, 14 (2011).

27. Toledano, M. B. *et al.* An international prospective cohort study of mobile phone users and health (COSMOS): factors affecting validity of self-reported mobile phone use. *Int. J. Hyg. Environ. Health* (2017) doi:https://doi.org/10.1016/j.ijheh.2017.09.008.

28. Tokola, K., Kurttio, P., Salminen, T. & Auvinen, A. Reducing overestimation in reported mobile phone use associated with epidemiological studies. *Bioelectromagnetics* **29**, 559–563 (2008).

29. Aydin, D. *et al.* Impact of random and systematic recall errors and selection bias in case-control studies on mobile phone use and brain tumors in adolescents (CEFALO study). *Bioelectromagnetics* **32**, 396–407 (2011).

30. Vrijheid, M., Deltour, I., Krewski, D., Sanchez, M. & Cardis, E. The effects of recall errors and of selection bias in epidemiologic studies of mobile phone use and cancer risk. *J. Expo. Sci. Environ. Epidemiol.* **16**, 371–84 (2006).

31. Tettamanti, G. *et al.* Long-term effect of mobile phone use on sleep quality: Results from the cohort study of mobile phone use and health (COSMOS). *Environ. Int.* **140**, (2020).

32. Auvinen, A. *et al.* Headache, tinnitus and hearing loss in the international Cohort Study of Mobile Phone Use and Health (COSMOS) in Sweden and Finland. *Int. J. Epidemiol.* **48**, 1567–1579 (2019).

33. Carter, B., Rees, P., Hale, L., Bhattacharjee, D. & Paradkar, M. S. Association between portable screen-based media device access or use and sleep outcomes a systematic review and meta-analysis. *JAMA Pediatr.* **170**, 1202–1208 (2016).

34. Fyhri, A. & Klæboe, R. Road traffic noise, sensitivity, annoyance and self-reported health-A structural equation model exercise. *Environ. Int.* **35**, 91–97 (2009).

35. Zijlema, W. L., Morley, D. W., Stolk, R. P. & Rosmalen, J. G. M. Noise and somatic symptoms: A role for personality traits? *Int. J. Hyg. Environ. Health* **218**, 543–549 (2015).

36. Héritier, H. *et al.* The Association between Road Traffic Noise Exposure, Annoyance and Health-Related Quality of Life (HRQOL). *Int. J. Environ. Res. Public Health* **11**, 12652–12667 (2014).

37. Baliatsas, C., van Kamp, I., Hooiveld, M., Lebret, E. & Yzermans, J. The relationship of modern health worries to non-specific physical symptoms and perceived environmental sensitivity: A study combining self-reported and general practice data. *J. Psychosom. Res.* **79**, 355–361 (2015).

38. Martens, A. L. *et al.* Longitudinal associations between risk appraisal of base stations for mobile phones, radio or television and non-specific symptoms. *J. Psychosom. Res.* **112**, 81–89 (2018).

39. Martens, A. L. *et al.* Modeled and perceived RF-EMF, noise and air pollution and symptoms in a population cohort. Is perception key in predicting symptoms? *Sci. Total Environ.* **639**, 75–83 (2018).

40. Tsarna, E. *et al.* Associations of Maternal Cell-Phone Use during Pregnancy with Pregnancy Duration and Fetal Growth in 4 Birth Cohorts. *Am. J. Epidemiol.* **188**, 1270–1280 (2019).

41. Uijl, A. *Opportunities and challenges of real-world data in heart failure*. (2020).

Appendices

ENGLISH SUMMARY

Everybody is exposed to radiofrequency electromagnetic fields (RF-EMF) in daily life, either by their own communications device(s) or by mobile phone base stations or radio transmitters in their living environment. Although a lot of experimental and epidemiological research has been conducted there are still concerns about the potential health effects of RF-EMF among scientists and in society. Uncertainties remain in RF-EMF epidemiology, such as scarcity of longitudinal epidemiological studies, lack of a reliable exposure assessment methods that can be applied in large observational epidemiological studies and absence of a clear biological mechanism. Furthermore for symptoms the role of perceived exposure in combination with the role of objectively assessed exposure is not clear. This thesis addresses some of these uncertainties in RF-EMF epidemiology such as investigating the repeatability of self-reported use of mobile communication devices and improvement of exposure assessment of mobile phone use by combining self-reported and operator-recorded data. Furthermore we examined the role of perception besides modeled exposure for health symptoms.

The first part of my thesis was about improving RF-EMF epidemiology by harmonizing and combining data from multiple prospective cohort studies and by evaluating and determining exposure assessment of RF-EMF by improving the accuracy of (self-reported) mobile phone use. Chapter two provides an overview of the data collected within the LIFEWORK cohort, the Dutch contribution to the international COSMOS study, in which we combined data from three prospective cohorts: Nightingale, EPIC-NL and AMIGO. Besides displaying baseline characteristics and environmental exposure distributions we also investigated repeatability of the baseline questionnaire in a subset of participants which filled in the questionnaire twice. The application of statistical approaches to improve exposure assessment of mobile phone use for epidemiological studies in the prospective COSMOS study was evaluated in chapter three. In this chapter we combined both self-reported and operator-recorded mobile phone data by several statistical methods e.g. regression calibration, complete-case and multiple imputation. Our study showed that combing different sources (self-reported and operator recorded) by regression calibration resulted in limited bias and outperformed other statistical methods such as complete-case and multiple imputation methods and could be used in future epidemiological investigations.

In chapter four the case ascertainment method we developed to identify Parkinson's Disease (PD) within two large prospective cohort studies (AMIGO and EPIC-NL) is described. We included multiple information sources like registry data (mortality, electronic medical records) and self-reported data. We evaluated our developed likelihood score, existing of four categories, with known (strong) etiological risk factors of Parkinson's Disease. Within EPIC-NL, risk analyses for the highest probability category showed strong associations with known risk factors of PD. Furthermore we compared the likelihood score with general practitioner information. Our study showed that especially mortality

records, electronic medical records (ICPC codes) and self-reported PD diagnosis and PD medication, were valid sources for PD case ascertainment in our study population. For epidemiological studies for which no specific registry is available, as is the case for Parkinson's Disease, it is important to combine multiple information sources to maximize sensitivity and specificity in the disease outcome measure.

In the last part of this thesis we expanded our knowledge on possible health effects of RF-EMF. Results in chapter five showed associations between birth outcomes and maternal self-reported mobile phone use in four birth cohorts. We examined several birth outcomes including pregnancy duration, birth weight and fetal growth. Our study found no effects with birth weight or fetal growth. Maternal mobile phone use during pregnancy appeared to be associated with shorter pregnancy duration and increased risk of preterm birth. Results were rather stable among three of the birth cohorts but we found a reversed association in one birth cohort. We could not adjust for maternal stress in this study, although it has been shown to be an important confounder.

In chapter six we investigated the impact of both modeled and perceived exposure to RF-EMF from mobile phone base stations, road traffic noise and air pollution on self-reported health outcomes such as, non-specific symptoms, sleep disturbance and respiratory symptoms. We used longitudinal data from the Dutch AMIGO cohort. Modeled RF-EMF exposure from mobile phone base stations was not associated with self-reported symptoms. Modeled air pollutants and noise exposure from road traffic were associated with increased symptoms scores. Perceived exposures showed overall stronger associations with self-reported health outcomes than modeled exposures. Risk estimates were stronger for non-specific symptom scores than for more specific health outcomes such as sleep disturbance and respiratory symptoms. When both modeled and perceived exposure was included as predictors in a single model, the impact of modeled exposure on self-reported health outcomes disappeared or strongly diminished. This shows the importance of perceived exposures in certain health effects which is a relatively under explored area in environmental epidemiology.

In chapter seven the findings are placed in a broader context within the RF-EMF epidemiology field. Furthermore in this last chapter the opportunities and challenges of using multiple information sources for assessment of exposure and health outcomes are discussed. This thesis showed that using multiple information sources for assessment of both exposure and health outcomes within observational epidemiological studies is feasible. The findings presented in this thesis will be input for future epidemiological analyses within the COSMOS study and for studies on Parkinson's Disease and environmental exposures within the AMIGO and EPIC-NL cohorts. The methods proposed in this thesis could also be applied in other (environmental) epidemiological studies. Further integration of multiple sources of RF-EMF by integrative exposure assessment models in epidemiological studies is essential to get more definitive answers on potential effects of RF-EMF exposure.

NEDERLANDSE SAMENVATTING

Iedereen wordt in het dagelijks leven blootgesteld aan radiofrequente elektromagnetische velden (RF-EMV), hetzij door eigen communicatieapparaat(en), hetzij door zendmasten voor mobiele telefoons of radiozenders in de leefomgeving. Ondanks dat al veel epidemiologisch onderzoek is gedaan, waarbij geen eenduidige effecten zijn gevonden, zijn er nog steeds zorgen over de mogelijke gezondheidseffecten van RF-EMV bij wetenschappers en de samenleving. Er blijven onzekerheden bestaan in de mogelijke gezondheidseffecten van RF-EMV door het gebrek aan longitudinale onderzoeken, waardoor het moeilijk is om causaliteit te bepalen, het ontbreken van betrouwbare methoden voor de karakterisering van blootstelling aan RF-EMV die kunnen worden gebruikt in grote epidemiologische studies en de afwezigheid van een duidelijk biologisch mechanisme. Verder is de rol van perceptie, in combinatie met de rol van de daadwerkelijke blootstelling op gezondheidsklachten nog onduidelijk. Dit proefschrift behandelt enkele van de genoemde onzekerheden in de RF-EMV epidemiologie, zoals het onderzoeken van de herhaalbaarheid van zelf-gerapporteerd gebruik van communicatiemiddelen en het verbeteren van het karakteriseren van de blootstelling door mobiel telefoon gebruik middels zelf-gerapporteerde en door de mobiele telefoon provider verzamelde gegevens te combineren. Daarnaast is de rol van perceptie naast de daadwerkelijke blootstelling onderzocht.

Het eerste deel van mijn proefschrift gaat over het verbeteren van de RF-EMV epidemiologie door het combineren van gegevens uit meerdere prospectieve cohort onderzoeken en door het evalueren en bepalen van de karakterisering van de blootstelling aan RF-EMV door zowel te kijken naar de validiteit als naar de betrouwbaarheid van zelf-gerapporteerd mobiel telefoon gebruik. Hoofdstuk 2 geeft een overzicht van de gegevens die zijn verzameld binnen het LIFEWORK cohort. Het LIFEWORK cohort is de Nederlandse bijdrage aan de internationale COSMOS studie, waarvoor gegevens van drie prospectieve cohorten zijn gecombineerd, namelijk: Nightingale, EPIC-NL en AMIGO. Naast het weergeven van kenmerken van de studie populatie en verdelingen van verschillende milieublootstellingen is ook de repliceerbaarheid van de vragenlijst onderzocht.

De toepassing van verschillende statistische methoden om de blootstellingsbeoordeling van het gebruik van mobiele telefoons te verbeteren voor epidemiologische studies werd geëvalueerd in hoofdstuk drie. In dit hoofdstuk hebben we zowel zelf-gerapporteerde als door de provider verzamelde mobiele telefoongegevens gecombineerd doormiddel van verschillende statistische methoden, zoals regressiekalibratie, complete-case analyse en imputatie. Onze studie toonde aan dat het combineren van verschillende bronnen (provider en zelf-gerapporteerde data) door middel van regressiekalibratie slechts tot een beperkte vertekening leidde en beter presteerde dan de andere onderzochte statistische methoden, zoals complete-case en imputatie methoden. Regressiekalibratie kan dus worden gebruikt in toekomstig RF-EMV epidemiologisch onderzoek.

In hoofdstuk vier is een methode beschreven, die we ontwikkelden om de ziekte van Parkinson te identificeren binnen twee grote prospectieve onderzoeken (AMIGO en EPIC-NL). In deze methode zijn meerdere informatiebronnen samengevoegd, waaronder registratiegegevens (sterfte, elektronische medische dossiers) en zelf-gerapporteerde gegevens. We evalueerden de ontwikkelde score, bestaande uit vier categorieën, met bekende (sterke) etiologische risicofactoren voor de ziekte van Parkinson. De resultaten voor de hoogste categorie waren in lijn met de bestaande literatuur voor EPIC-NL. Verder vergeleken we de score met informatie verkregen bij huisartsen. Onze studie toonde aan dat met name (een combinatie van) gegevens uit sterfteregistratie, gegevens uit elektronische medische dossiers (ICPC-codes) en zelf-gerapporteerde PD-diagnose en PD-medicatie, valide bronnen zijn om de ziekte van Parkinson te identificeren. Voor epidemiologisch onderzoek naar gezondheidsuitkomsten waarvoor geen specifieke registratie beschikbaar is, zoals voor de ziekte van Parkinson, blijft het belangrijk om informatie uit meerdere bronnen te combineren.

In het laatste deel van dit proefschrift hebben we onze kennis over mogelijke gezondheidseffecten van RF-EMV uitgebreid. Hoofdstuk vijf laat de associaties zien tussen geboorte-uitkomsten en zelf-gerapporteerd mobiel telefoongebruik door de moeder tijdens de zwangerschap in vier verschillende geboortecohorten. We onderzochten verschillende geboorte-uitkomsten, waaronder zwangerschapsduur, geboortegewicht en foetale groei. Onze studie vond geen effect van mobiel telefoongebruik van de moeder op geboortegewicht en foetale groei. Het gebruik van mobiele telefoons door de moeder tijdens de zwangerschap kon wel in verband worden gebracht met een kortere zwangerschapsduur en een verhoogd risico op vroeggeboorte. De resultaten waren vrij vergelijkbaar binnen de vier geboortecohorten, echter binnen één cohort vonden we een omgekeerd verband. In alle geboortecohorten hebben we niet gecorrigeerd voor stress, terwijl dit een belangrijke factor is die de gevonden resultaten kan beïnvloeden.

In hoofdstuk zes onderzochten we de impact van zowel gemodelleerde als ervaren blootstelling aan RF-EMV van zendmasten voor mobiele telefoons, geluid van wegverkeer en luchtvervuiling op zelf-gerapporteerde gezondheidsklachten, zoals niet-specifieke klachten, slaapverstoring en respiratoire symptomen. We gebruikten longitudinale gegevens van het Nederlandse AMIGO cohort. Gemodelleerde RF-EMV blootstelling van zendmasten voor mobiele telefoons bleek niet geassocieerd met zelf-gerapporteerde symptomen. Gemodelleerde blootstelling aan luchtvervuiling en geluid van wegverkeer waren wel geassocieerd met verhoogde gezondheidsklachten. Ervaren blootstellingen vertoonden over het algemeen sterkere associaties met zelf-gerapporteerde gezondheidsklachten dan gemodelleerde blootstellingen. Effecten waren relatief groter voor niet-specifieke klachten dan voor specifieke gezondheidsuitkomsten zoals slaapverstoring en respiratoire symptomen. Wanneer zowel de gemodelleerde als de bijbehorende ervaren blootstelling gecombineerd werden geanalyseerd, nam de impact van gemodelleerde blootstelling op zelf-gerapporteerde gezondheidsklachten sterk

af of verdween. Alleen het effect van de ervaren blootstelling op gezondheidsklachten was dan significant. Dit toont het belang aan van ervaren blootstelling voor bepaalde gezondheidseffecten in epidemiologische onderzoeken wat een relatief onderbelicht onderwerp is in de milieuepidemiologie.

In hoofdstuk zeven worden de bevindingen in een bredere context geplaatst binnen het veld van de RF-EMV epidemiologie. Verder worden in dit laatste hoofdstuk kansen en uitdagingen van het gebruik van meerdere informatie bronnen bij de karakterisering van blootstelling en gezondheidseffecten besproken. Het werk in dit proefschrift toonde aan dat het gebruik van meerdere informatie bronnen voor karakterisering van blootstelling en gezondheidseffecten binnen observationele epidemiologische studies haalbaar en wenselijk is. De onderzoeken in dit proefschrift zullen input zijn voor toekomstig epidemiologisch onderzoeken. Zo zullen de regressie kalibratie gegevens uit hoofdstuk drie gebruikt worden in de COSMOS studie. Binnen EPIC-NL en AMIGO zal onderzoek naar de ziekte van Parkinson in relatie tot verschillende omgevingsblootstellingen gedaan worden. De methoden die in dit proefschrift worden voorgesteld kunnen ook worden toegepast voor andere blootstellingen en gezondheidseffecten. Verdere integratie van meerdere bronnen van RF-EMV door middel van integratieve blootstellingsmodellen in epidemiologische onderzoeken is essentieel om meer definitieve antwoorden te krijgen over de mogelijke effecten van blootstelling aan RF-EMV.

AFFILIATIONS OF CONTRIBUTORS

Virissa Lenters, Pauline Slottje, Hans Kromhout, Roel Vermeulen, Inka Pieterson, Ermioni Tsarna, Anke Huss, Astrid L. Martens, Maciej Strak, Lützen Portengen, Eugenio Traini
Division of Environmental Epidemiology, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands

Petra H Peeters, Roel Vermeulen, Roel, W.M. Monique Verschuren
Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

Pauline Slottje, Astrid L. Martens, Tjabe Smid, Danielle Timmermans
Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands

Anouk Pijpe, Matti A Rookus, Flora E. van Leeuwen
Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands

Roel Vermeulen, Bas Bueno-de-Mesquita
Imperial College London, Department of Biostatistics and Epidemiology, London, United Kingdom

Joke C. Korevaar, Robert A.Verheij
Netherlands Institute for Health Services Research (NIVEL), Utrecht, The Netherlands

W.M. Monique Verschuren, Bas Bueno-de-Mesquita, Danielle Timmermans, Wim Swart
National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Laura Ellen Birks, Mònica Guxens, Martine Vrijheid, Elisabeth Cardis
ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain

Laura Ellen Birks, Mònica Guxens, Martine Vrijheid, Elisabeth Cardis
Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona, Spain

Laura Ellen Birks, Mònica Guxens, Martine Vrijheid, Elisabeth Cardis, Ferran Ballester, Aitana Lertxundi
Spanish Consortium for Research on Epidemiology and Public Health, Instituto de Salud Carlos III, Madrid, Spain

Mònica Guxens
Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Centre Sophia Children's Hospital, Rotterdam, the Netherlands

Ferran Ballester, Llúcia González Safont
Epidemiology and Environmental Health Joint Research Unit, Foundation for the Promotion of Health and Biomedical Research of Valencia Region, Universitat Jaume I–Universitat de València, Valencia, Spain

Mina Ha, Hyung-Ryul Lim
Department of Preventive Medicine, College of Medicine, Dankook University, Cheonan, South Korea

Ana Jiménez-Zabala, Aitana Lertxundi
Environment Epidemiology and Child Development Area, BIODONOSTIA Health Research Institute, San Sebastian, Spain

Ana Jiménez-Zabala
Public Health Division of Gipuzkoa, Basque Government, San Sebastian, Spain

Leeka Kheifets, Madhuri Sudan
Department of Epidemiology, School of Public Health, University of California, Los Angeles, Los Angeles, California

Aitana Lertxundi
Department of Preventive Medicine and Public Health, Universidad del Pais Vasco/Euskal Herriko Unibertsitatea, Leioa, Spain

Jorn Olsen, Madhuri Sudan
Department of Clinical Epidemiology, Aarhus University, Aarhus, Denmark

Llúcia González Safont
Unitat Predepartamental de Medicina, Universitat Jaume I, Castelló de la Plana, Spain

Tanja Vrijkotte
Department of Public Health, Amsterdam Public Research Institute, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands

Madhuri Sudan
College of Osteopathic Medicine of the Pacific, Western University of Health Sciences, Pomona, California

Anssi Auvinen, Katja Kojo, Sirpa Heinävaara
STUK – Radiation and Nuclear Safety Authority, Environmental Surveillance, Helsinki, Finland

Anssi Auvinen, Katja Kojo, Sirpa Heinävaara

University of Tampere, School of Health Sciences, Tampere, Finland

Giorgio Tettemanti, Anders Ahlbom, Lena Hillert, Maria Feychting

Karolinska Institute, Institute of Environmental Medicine, Stockholm, Sweden

Paul Elliott, Mireille Toledano

MRC-PHE Centre for Environment and Health, Imperial College London, School of Public Health, Department of Epidemiology and Biostatistics, London, UK

Joachim Schüz, Isabelle Deltours

International Agency for Research on Cancer (IARC), Section of Environment and Radiation, Lyon, France

Joachim Schüz, Aslak Harbo Poulsen

Danish Cancer Society Research Center, Copenhagen, Denmark

LIST OF PUBLICATIONS

This thesis

Reedijk M*, Lenters V*, Slottje P, et al. Cohort profile : LIFEWORK , a prospective cohort study on occupational and environmental risk factors and health in the Netherlands. Published Online First: 2018. doi:10.1136/bmjopen-2017-018504

Reedijk M*, Martens AL*, Smid T, et al. Modeled and perceived RF-EMF, noise and air pollution and symptoms in a population cohort. Is perception key in predicting symptoms? Sci Total Environ 2018;639:75–83. doi:10.1016/j.scitotenv.2018.05.007

Tsarna E, **Reedijk M,** Birks LE, et al. Associations of Maternal Cell-Phone Use during Pregnancy with Pregnancy Duration and Fetal Growth in 4 Birth Cohorts. Am J Epidemiol 2019;188:1270–80. doi:10.1093/aje/kwz092

Reedijk M, Huss A, Verheij RA, et al. Parkinson's disease case ascertainment in prospective cohort studies through combining multiple health information resources. PLOS ONE 2020*;* 15(7): e0234845. https://doi.org/10.1371/journal.pone.0234845

Reedijk M*, Portengen L*, Auvinen A, et al. Regression calibration of self-reported mobile phone use to improve quantitative risk estimation in the COSMOS study, in preparation

Other publications

Reedijk M, Zock JP, Janssen NAH, Verkenning haalbaarheid gezondheidsonderzoek werknemers Schiphol, RIVM-briefrapport 2021-0094.

Reedijk M, Hin J, van Kamp I, Factsheet gezondheidseffecten van windturbinegeluid, juli 2021, RIVM

Martens, A. L., Slottje, P., Timmermans, D. R. M., Kromhout, H., **Reedijk, M**., Vermeulen, R. C. H., &Smid, T. (2017). Modeled and Perceived Exposure to Radio-Frequency Electromagnetic Fields From Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort. American Journal of Epidemiology, 1–10. https://doi.org/10.1093/aje/kwx041

Martens, A. L., Slottje, P., Timmermans, D. R. M., Kromhout, H., **Reedijk, M**., Smid, T., Vermeulen, R. C. H. (2017). Reply to letter to the editor by Dr. Mortazavi on our article "Modeled and Perceived Exposure to Radio-Frequency Electromagnetic Fields from Mobile-Phone Base Stations and the Development of Symptoms Over Time in a General Population Cohort". American Journal of Epidemiology.

Reedijk M, Van Leuken JPG, Van Der Hoek W. Particulate matter strongly associated with human Q fever in The Netherlands: an ecological study. *Epidemiol Infect* 2013;:1–11.http://www.ncbi.nlm.nih.gov/pubmed/23481147

ABOUT THE AUTHOR

Marije Reedijk was born on 11 March 1988 in Utrecht. In 2009 she completed her Bachelor's in Health Sciences at VU Amsterdam. In 2009 she started a second Bachelor's in Social Geography and Planning at Utrecht University which she completed in 2011. She combined these two research areas in the Master Health Sciences with a specialization in Infectious Disease and Health. The research for her Master's Degree was conducted at the National Institute for Public Health and Environment (RIVM). Here, she studied environmental risk factors of Q fever in the Netherlands under supervision of Wim van der Hoek. After graduation with distinction in 2012 she continued working at the RIVM as a junior researcher on the topic environmental risk factors and Q fever. In 2012 and 2013 she travelled by bicycle through South and Central America. In 2014 she started with her PhD project at the Institute for Risk Assessment Sciences and Julius Center for Health Sciences an Primary Care under the supervision of Roel Vermeulen, Hans Kromhout and Petra Peeters. The results of this PhD project are presented in this thesis. Her project was linked to interdisciplinary collaborations with Dutch and European projects, including COSMOS, AMIGO, Nightingale and EPIC-NL. During her PhD she followed the postgraduate Master Epidemiology, specialization Environmental and Occupational Epidemiology from Utrecht University, which she completed in 2018. In 2018 she completed the young adult education in theme centered interaction (TCI). In 2020 she started working at the National Institute for Public Health and Environment as a researcher.

DANKWOORD

Het promotie avontuur zit er bijna op, het boekje is met het schrijven van dit dankwoord echt af. Een nieuwe baan, een mooie reis, twee schatten van kinderen, corona, motivatie problemen, verbouwingen, fietsen en 101 andere leuke dingen zorgde voor de nodige vertraging van het afronden. Maar ik kan nu na bijna 8 jaar eindelijke op de vraag “hoe gaat het met je proefschrift” antwoorden: **het boekje is af**!

Ook ik had mijn proefschrift niet kunnen schrijven zonder de hulp van anderen en deze mensen wil ik hier dus graag bedanken.

Een speciale dank gaat uit naar mijn promotoren: Roel, Hans en Petra. Bedankt voor deze mooie kans die jullie me hebben geboden. Bedankt voor jullie begeleiding de afgelopen jaren en jullie kritische blik op mijn stukken, ik heb veel van jullie geleerd over epidemiologie en onderzoek doen. Roel, het was soms moeilijk een plekje te vinden in je agenda maar met de weinige tijd die je beschikbaar had bracht je me altijd weer een stapje verder en waren de hobbels die ik zag opeens niet meer zo hoog. Hans, bij jou kon ik altijd binnenlopen om een vraag te stellen over EMV blootstelling of over een verhaal over fietsen. Petra, bedankt dat ik in de eerste jaren kon leren van jou epidemiologische kennis en ervaring met vragenlijstonderzoek, jij zorgde ook altijd voor een realistische planning.

Samenwerken doe ik graag, het is onder andere te zien aan het aantal co-auteurs op mijn artikelen. Door jullie feedback zijn het nog mooiere eindproducten geworden. Bedankt Astrid, voor de samenwerking aan ons paper die we tijdens onze zwangerschappen samen hebben geschreven. Het was een plezier om met je samen te werken met jouw psychologische kennis! Virissa bedankt voor je begeleiding in de beginfase van mijn promotie avontuur en de input op het LIFEWORK artikel. Lützen, bedankt voor het delen van jouw statistische kennis en het doen van de analyses voor het COSMOS paper. Luuk en Anke bij jullie kon ik altijd terecht voor kennis op het gebied van EMV, het waren altijd plezierige gesprekken en brachten me vaak weer een stapje verder. Inka, jij wist altijd precies te vertellen hoe het zat met de AMIGO data en ik kon altijd binnen lopen voor een praatje of om te klagen als iets niet werkte. Eef, bedankt voor het maken van de website tijdens het veldwerk voor AMIGO. Ingrid, bij jou kon ik altijd terecht voor praktische tips en bedankt voor alle belletjes voor het Parkinson project. Joke, bedankt voor je hulp bij het veldwerk van EPIC-NL. Guido, Ermina, Jara en Leandro, bedankt voor jullie hulp met de dataverzameling en analyse tijdens jullie masterstage, het was een plezier om jullie te mogen begeleiden. Dear COSMOS team, thank you for sharing the data and knowledge on EMF, it was a pleasure working together on this important project. Ook de leden voor de beoordelingscommissie, prof. dr. Lidwien Smit, prof. dr. ir. Erik Lebret, prof. dr. Chantal Kemner, prof. dr. Gerard van Rhoon, dr. Irene van Kamp, dr. Miriam Sturkenboom, wil ik bedanken voor het lezen en beoordelen van mijn proefschrift.

Het IRAS was een fijne plek om te werken. Dus ik wil bij deze alle collega's bedanken. Na de lunchwandelingen, koffiepauzes en seminars had ik altijd weer nieuwe energie en wetenschappelijke input om verder te gaan. Maar ook de uitjes na het werk zoals het Smartlappen en Herfstbok festival, carnaval in Maastricht, spelletjesavonden, vrijdagmiddagborrels en natuurlijk de ski trip met Floor en José, zorgde er voor dat ik altijd met plezier naar mijn werk ging. Lieve 3.19B kamergenoten: Aileen, Joseph, Calvin, Edith, Jules, Daniel en Ilse. Ik voelde me altijd thuis bij jullie op de kamer. Zowel tijdens de koffie (of in mijn geval meestal thee) momentjes in onze kamer als tijdens de leuke uitjes naar FC Utrecht, de Keukenhof, internationale etentjes en alcoholproeverijen. Jules en Luuk, ik stond naast jullie als paranimf en ben blij dat jullie mij willen vergezellen tijdens de verdediging! Ook mijn collega's van het Juliuscentrum wil ik bedanken, de onco-overleggen in het begin van mijn promotietraject zorgde altijd weer voor een andere wetenschappelijke blik op mijn projecten. Ook mijn RIVM collega's wil ik bedanken tijdens de laatste loodjes van het afronden van mijn proefschrift en de interessante projecten waar we aan werken.

Tot slot wil ik ook graag mijn familie en vrienden bedanken, jullie zorgde naast mijn werk voor de ontspanning en jullie luisterde altijd naar mijn verhalen over mijn onderzoek. Lieve opa's en oma's bedankt voor al het oppassen de afgelopen 4 jaar en jullie luisterende oor. Zonder jullie was het boekje nu nog niet af geweest. Mama, helaas kan je mijn boekje niet lezen, je steunde me altijd, las al mijn stukken door op de middelbare school en ik weet dat je trots zou zijn geweest. Wat had ik je graag in de zaal gehad ♥. Maurice, bedankt voor je betrokkenheid en interesse, ik ben heel blij met jou als broertje. Nikki en Rosie, door jullie kwam mijn proefschrift op de tweede plaats en duurde het iets langer voordat ik klaar was, maar jullie vrolijkheid zou ik nooit willen missen. Menno, bedankt dat je er altijd voor me bent en voor de nodige afleiding tijdens mijn promotietraject. Je nam me mee naar het strand of bos als ik het even niet meer zag zitten zodat ik daarna weer met nieuwe energie verder kon. Wat hebben we een heerlijk leven met zijn viertjes, ik kijk uit naar ons kampeeravontuur in Amerika.

De Bilt, 20 februari 2022

Institute for Risk Assessment Sciences, Utrecht University