An international prospective cohort study of mobile phone users and health (COSMOS): factors affecting validity of self-reported mobile phone use

Authors: Mireille B. Toledano PhD[a,b*], Anssi Auvinen PhD[c,d], Giorgio Tettamanti PhD[e], Yang Cao PhD[e,f], Maria Feychting PhD[e], Anders Ahlbom PhD[e], Karin Fremling MSc[e], Sirpa Heinävaara PhD[g], Katja Kojo PhD[d], Gemma Knowles PhD[a,b], Rachel B. Smith PhD[a,b], Joachim Schüz PhD[h], Christoffer Johansen PhD[i,j], Aslak Harbo Poulsen PhD[j], Isabelle Deltour PhD[h], Roel Vermeulen PhD[k], Hans Kromhout PhD[k], Paul Elliott PhD[a,b], Lena Hillert PhD[e,l]

Affiliations:

[a]MRC-PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, Norfolk Place, London, W2 1PG, UK; [b]National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Health Impact of Environmental Hazards, School of Public Health, Imperial College London, Norfolk Place, London, W2 1PG, UK. [c]School of Health Sciences, University of Tampere, FI-33014, Tampere, Finland; [d]Radiation and Nuclear Safety Authority (STUK), 00811 Helsinki, Finland; [e]Institute of Environmental Medicine, Karolinska Institutet, SE-171 77 Stockholm, Sweden; [f]Clinical Epidemiology and Biostatistics, School of Medical Sciences, Örebro University, SE-701 82 Örebro, Sweden: [g]Finnish Cancer Registry, Mass Screening Registry, Unioninkatu 22, FI-00130 Helsinki, Finland; [h]International Agency for Research on Cancer (IARC), Section of Environment and Radiation, 69372 Lyon, France; [i]Oncology clinic, Finsen Center, Copenhagen, Denmark; [j]The Danish Cancer Society Research Center, DK-2100 Copenhagen, Denmark; [k]Institute for Risk Assessment Sciences, Utrecht University, 3508 TD Utrecht, The Netherlands; [l]Centre

for Occupational and Environmental Medicine, Stockholm County Council, 104 22 Stockholm, Sweden.

Corresponding author: Dr Mireille B Toledano, MRC-PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, St Mary's Campus, Norfolk Place, London, W2 1PG, UK, Tel: +44 (0)2075943298, Fax: +44 (0)2075940768, E-mail: m.toledano@imperial.ac.uk

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 at King's College London and Imperial College London in partnership with 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 number 2013-CRD-17 and the International Agency for Research on Cancer.

Abstract

This study investigates validity of self-reported mobile phone use in a subset of 75 993 adults from the COSMOS cohort study. Agreement between self-reported and operator-derived mobile call frequency and duration for a 3-month period was assessed using Cohen's weighted Kappa (κ). Sensitivity and specificity of both self-reported high ($\geq$10 calls/day or $\geq$4 hours/week) and low ($\leq$6 calls/week or <30 minutes/week) mobile phone use were calculated, as compared to operator data. For users of one mobile phone, agreement was fair for call frequency (κ=0.35, 95% CI: 0.35, 0.36) and moderate for call duration (κ =0.50, 95% CI: 0.49, 0.50). Self-reported low call frequency and duration demonstrated high sensitivity (87% and 76% respectively), but for high call frequency and duration sensitivity was lower (38% and 56% respectively), reflecting a tendency for greater underestimation than overestimation. Validity of self-reported mobile phone use was lower in women, younger age groups and those reporting symptoms during/shortly after using a mobile phone. This study highlights the ongoing value of using self-report data to measure mobile phone use. Furthermore, compared to continuous scale estimates used by previous studies, categorical response options used in COSMOS appear to improve validity considerably, most likely by preventing unrealistically high estimates from being reported.

Keywords

Cellular phone, telecommunications, radiofrequency, electromagnetic fields, validation

Introduction

The possible adverse health effects of radiofrequency exposure from mobile phones are of considerable public and scientific interest. Overall, the balance of evidence does not suggest an excess risk, with studies on mobile phone use and cancer, primarily brain tumours, mostly reporting risk estimates close to unity,(1-8) though some have reported increased risk of brain tumours among the heaviest mobile phone users when considering long-term (>10 years) use.(8-12) However, the majority of these cancer studies are limited by their reliance on subjective, self-reported measures of telephone use in the past (4, 8-16) which are prone to substantial recall error,(17-22) and are case-control studies (8, 10-12, 14) which are also prone to recall and selection bias.(23, 24) Evidence for potential effects of mobile phone use on other health outcomes (e.g. headaches, migraines, fatigue, cognition, sleep disturbance, dizziness, hearing loss, etc) is largely based on cross-sectional studies, with inconsistent results.(1, 25-29)

Non-differential random error in continuous exposure measures is more likely, but not guaranteed, to bias risk estimates towards the null,(30) whereas the impact of non-differential misclassification of categorical measures, (31, 32) and systematic and differential error is less predictable, and can attenuate, strengthen, or reverse a true association, or produce spurious associations.(30, 33-36) Non-differential random error or misclassification also reduces statistical power to detect a true association.(30)

Previous validation studies have generally reported fair-to-moderate agreement between self-reported mobile phone use and mobile network operator data,(17, 18, 20-22, 37-41) and have consistently demonstrated substantial overestimation of call duration by self-reported measures,(18, 21, 38, 40-43) particularly among the heaviest users.(38) Conversely, call frequency tends to be slightly underestimated by self-reported measures,(21, 38, 41) although some studies report overestimation for both frequency and duration.(17, 40) However, these findings are often based on small numbers [e.g. n<100 (17, 21, 41, 42)], and some are drawn from case-control studies of mobile phone use and risk of cancer,(21, 22, 38, 43) thus limiting generalizability to the general population. Moreover, it is unknown if validity differs between subgroups of the population e.g. between males and females, different age groups, users of more than one mobile phone, those experiencing symptoms when using a mobile phone, or those concerned about mobile phones/base stations and health. For such groups, both level of mobile phone use and accuracy of self-reporting may be associated, potentially resulting in differential error according to usage.

This study investigates the validity of self-reported mobile phone use, by comparing cross-sectional baseline data on self-reported and operator-derived mobile phone use (frequency and duration of calls), in a large sub-population of 75 993 adults participating in the COSMOS (COhort Study of MObile phone uSe and health) project. It also investigates, for the first time, validity among general population subgroups, e.g. those who experience symptoms during mobile phone use or have concerns related to mobile phones.

Participants and methods

Sampling and participants

The study design for the international prospective cohort study COSMOS has been described in detail elsewhere.(44, 45) The target population for COSMOS was adult mobile phone users, aged 18-69 years, in 5 European countries: Denmark, Finland, the Netherlands, Sweden and the UK, and recently a 6th cohort has been initiated in France.

This analysis focuses on participants recruited into the study in Finland, Sweden and the UK between 2007 and 2010. Participants were identified by stratified random sampling (based on call time and age; in Finland and the UK also on sex) from subscriber lists of the major network operators in each country. Eligible for inclusion were those who gave permission for COSMOS to access their operator data and who answered the baseline questionnaire: comprising 13 070 participants in Finland, 50 736 participants in Sweden and 62 938 participants in the UK. We further limited the analysis to those who reported one or two mobile phone numbers (used in the last three months) which could each be matched to a single network operator (i.e. participants who switched operators within this time were excluded), and for which complete operator data were successfully obtained for the three months preceding the completion of the baseline questionnaire for these mobile phone numbers (N.B. not all mobile phone operators had been contacted at the time of compilation of data for these analyses). This left 75 993 participants (6 229, 30 874, and 38 890 from Finland, Sweden and the UK, respectively) in this analysis.

Consent and ethical approval

COSMOS was approved by the local research ethics committees in each country. Participants gave written or electronic informed consent.

Questionnaire data

The COSMOS baseline questionnaire was administered as a web-based survey (Finland and UK) and/or on paper (Finland and Sweden). It included questions on past and recent use of mobile phones, symptoms during mobile phone use, risk perception related to mobile phone use, and demographic information.(44, 45)

Self-reported mobile phone use

Participants were asked to report frequency and duration of mobile phone voice calls for the preceding three months via the following two questions:

"Over the last three months, how often did you talk on a mobile phone?" with the response options: <1 call per week (Finland and Sweden only; the UK web-based questionnaire filtered out these respondents in a previous question), 1-6 calls per week, 1-9 calls per day, ≥10 calls per day.

"Over the last three months, on average, how much time per week did you spend talking on a mobile phone?" with the response options: <5 minutes, 5-29 minutes, 30-59 minutes, 1-3 hours, 4-6 hours, >6 hours.

Questionnaire response category cut-point choices were informed by distributions observed in operator data in the COSMOS pilot study, and also in the Interphone study,(8, 44), in order to give categories that would be distinct based on those distributions, and would also appear logical to participants. In the UK questionnaire, the highest categories were expanded to reflect high and rapidly increasing mobile phone use in the general population (i.e. "≥10 calls per day" was expanded to "10-29 calls/day" and "≥30 calls/day", and ">6 hrs/week" was expanded to "7-9 hours/week" and "≥10 hours/week"). For this analysis, these categories were collapsed to be comparable with Finland and Sweden.

Respondents were asked to provide the phone numbers of the two (Finland and Sweden) or three (UK only) mobile phones they used most frequently, and to indicate the proportion of total calls made by the respondent on each phone, and the proportion of calls made by other people on each phone. In these analyses, the third phone reported by 0.3% of UK participants was ignored for comparability with the Swedish and Finnish data.

Symptoms during mobile phone use

Participants were asked if they experienced symptoms ("*no symptoms, headache, dizziness, numbness in hands, nausea, hearing loss, tinnitus/ringing sound in ear, warming sensation on face and/or ear"*) whilst using, or shortly after using, a mobile phone. If participants reported warming sensations only (a common occurrence likely due to heat generated by the phone battery) they were excluded from the analysis of symptoms. Those reporting any other symptoms were classified as 'yes' for experiencing symptoms related to mobile phone use, and were compared to those reporting no symptoms.

Risk perception

Participants were asked if they were concerned *("no concern, a little concern, moderate concern, high concern, extreme concern"*) that mobile phone use, proximity to mobile phone masts (base stations), or new technology might affect their health. For analysis, participants were categorised as "no concern" vs. "any concern" for each of mobile phone use, base stations and new technology.

Operator-derived mobile phone use

All major network operators (four in both Sweden and UK, and three in Finland) were asked to provide information on incoming and outgoing calls for at least a three month period for consenting participants. Network operators were requested to provide data for a time period which overlapped with self-report data, or as near as possible. The processes by which operator data were matched and acquired in the UK, Sweden and Finland are described elsewhere.(44, 46) For analysis, continuous operator data were categorised to match the response categories for self-reported call frequency and duration. Operator call duration values >3 and <3.5 were rounded down to the 1-3 hours/week category, and values ≥3.5 and <4 were rounded up to the 4-6 hours/week category.

Statistical analyses

The proportions of participants who under-, correctly, and over-estimated their mobile phone use, compared to their operator data were calculated. The proportion of participants classified in the same usage category for both self-report and operator data and Cohen's weighted Kappa, a measure of inter-rater agreement for categorical data,(47, 48) were used to assess concordance between self-reported and operator data. Kappa values are generally interpreted as: ≤0=poor, 0.01–0.20=slight, 0.21–0.40=fair, 0.41–0.60=moderate, 0.61–0.80=substantial, and >0.8=excellent.(48) Call frequency was defined as high if ≥10 calls/day (12% of respondents in the sample) and low if ≤6 calls/week (33% of respondents in the sample). Call duration was defined as high if ≥4 hours/week (19% of respondents in the sample) and low if <30 minutes/week (31% of respondents in the sample). These high/low categories for analysis were chosen in order to get contrasting categories, e.g. a category for high exposure with more likely high exposure compared to a wider category. Sensitivity and specificity for high (vs. not high) and low (vs. not low) call frequency and duration, and 95% confidence intervals (95% CIs) were also calculated, as compared to operator data. Analyses were conducted for the whole sample, and also stratified by country and number of phones used (henceforth, 'one phone users' and 'two (or more) phone users'). Additional subgroup comparisons (pre-specified, based on age group, sex, symptoms, and risk perception) were conducted among one phone users only. Sensitivity analyses were conducted restricted to Swedish and Finnish data excluding the following groups: those who reported <10% of total use for the first phone (n=2229); those who reported <40% of total use for the two phones (n=1803); and those who reported other people regularly using their phone(s) (n=1309). For the UK participants this information was not available.

Results

Participants' characteristics and mobile phone use

Among included participants, 68 087 (90%) reported using only one mobile phone and 7 906 (10%) reported using two (or more) mobile phones. According to operator data, the majority of participants spent at least 30 minutes per week on their mobile phone (74%) and/or made at least one call per day (89%) (Table 1). Approximately 20% of participants spent at least 4 hours per week on calls and/or made at least 10 calls per day, and were thus defined as having high mobile phone use (Table 1). Compared with those who used one phone, two (or more) phone users were more likely to be male and had higher average call duration and frequency (for both self-reported and operator data) (Table 1). Overall, 10 933 (14%) reported experiencing symptoms whilst (or shortly after) using a mobile phone and 45 012 (59%) reported some level of concern about mobile phones and health (ranging from a little concern (36%) up to extreme concern (1%)).

Table 1. Participant characteristics and mobile phone use

	Total (n = 75 993)	One phone users (n = 68 087)	Two (or more) phone users (n = 7906)
	n (%)	**n (%)**	**n (%)**
Sex			
Men	34 041 (45)	29 713 (44)	4328 (55)
Women	41 879 (55)	38 306 (56)	3573 (45)
Missing	73 (0)	68 (0)	5 (0)
Age group			
18-33 years	19 756 (26)	18 099 (27)	1657 (21)
34-49 years	21 727 (29)	18 969 (28)	2758 (35)
≥ 50 years	34 351 (45)	30 875 (45)	3476 (44)
Missing	159 (0)	144 (0)	15 (0)
Self-reported call duration			
<5 min/week	3121 (4)	2938 (4)	183 (2)
5 to <30 min/week	20 535 (27)	18 917 (28)	1618 (20)
30 to <60 min/week	16 057 (21)	14 563 (21)	1494 (19)
1 to 3hours/week	21 414 (28)	19 134 (28)	2280 (29)
4 to 6 hours/week	8604 (11)	7437 (11)	1167 (15)
>6 hours/week[2]	6126 (8)	4982 (7)	1144 (14)
Missing	136 (0)	116 (0)	20 (0)
Operator-derived call duration			
<5 min/week	3425 (5)	3318 (5)	107 (1)
5 to <30 min/week	16 076 (21)	15 050 (22)	1026 (13)
30 to <60 min/week	11 947 (16)	10 825 (16)	1122 (14)
1 to 3 hours/week	29 338 (39)	26 263 (39)	3075 (39)
4 to 6 hours/week	10 185 (13)	8809 (13)	1376 (17)
>6 hours/week	5022 (7)	3822 (6)	1200 (15)
Self-reported call frequency			
Less than 1 call per week[1]	487 (1)	444 (1)	43 (1)
1-6 calls/week	24 539 (32)	22 848 (34)	1691 (21)
1-9 calls/day	41 633 (55)	37 165 (55)	4468 (56)
≥10 calls/day[2]	9169 (12)	7490 (11)	1679 (21)
Missing	165 (0)	140 (0)	25 (0)
Operator-derived call frequency			
Less than 1 call per week	435 (1)	428 (1)	7 (0)
1-6 calls/week	7711 (10)	7356 (11)	355 (4)
1-9 calls/day	52 251 (69)	47 371 (70)	4880 (62)
≥10 calls/day	15 596 (20)	12 932 (19)	2664 (34)

[1] Finland and Sweden only. The UK questionnaire filtered out, in the previous question, those who reported <1 call per week. Note: percentages are rounded to the nearest integer so totals may not equal 100. [2] In the UK questionnaire, the highest self-report response categories for call duration (“7-9 hours/week” and “≥10 hours/week”) and call frequency (“10-29

calls/day" and "⩾30 calls/day"] were collapsed to >6 hours/week and ≥10 calls/day respectively to be comparable with Finland and Sweden for analysis.

Comparison of self-report and operator data

Agreement

We found that a considerable proportion of respondents misclassified their mobile phone use (approximately 60% and 40% for call duration and frequency, respectively) (Table 2, Supplementary Tables 1 & 2). Approximately a third of the participants underestimated their mobile phone call duration and frequency. The proportion of participants overestimating mobile phone use was much lower (23% for duration and 5% for call frequency among one-phone users) (Table 2). This pattern was similar among one- and two (or more)- phone users and across the countries.

Table 2. Percentage of participants who underestimated, correctly estimated and overestimated their mobile phone use, by country and number of mobile phones

	Call frequency[1]			Call duration[1]		
	Under-estimate	**Correct estimate**	**Over-estimate**	**Under-estimate**	**Correct estimate**	**Over-estimate**
One phone users						
Finland	26.3	69.2	4.5	42.0	43.2	14.7
Sweden	35.8	59.4	4.8	35.7	43.9	20.3
UK	36.5	58.1	5.4	31.1	43.3	25.6
All	35.4	59.5	5.1	33.8	43.5	22.7
Two (or more) phone users						
Finland	29.4	63.9	6.6	45.8	41.8	12.4
Sweden	36.0	55.5	8.4	41.0	37.0	22.0
UK	36.9	56.1	6.9	36.5	37.9	25.6
All	36.0	56.1	7.9	39.9	37.5	22.6

[1]Agreement (%) calculated based on 3 categories for call frequency and 6 categories for call duration.

Weighted Cohen's Kappa and sensitivity

Agreement between self-reported and operator data was moderate for call duration (κ=0.50, 95% CI: 0.49, 0.50 and κ=0.41, 95% CI: 0.39, 0.42 for one- and two (or more)- phone users, respectively) and fair for call frequency (κ=0.35, 95% CI: 0.35, 0.36 and κ=0.30, 95% CI: 0.28, 0.31 for one- and two (or more)- phone users, respectively) (Table 3). For one phone users, sensitivity of the self-report questionnaire was 87% and 76% for low call frequency and low call duration, respectively, and 38% and 56% for high call frequency and high call duration, respectively. Compared with one phone users, two (or more) phone users showed lower agreement between self-report and operator data, and lower sensitivity of self-report for low use (72% and 66% for low call frequency and low call duration respectively), but slightly greater sensitivity for high use (43% and 58% for high call frequency and high call duration respectively). Sensitivity of self-report for high call duration was greater for the UK compared with Finland and Sweden.

Table 3. Agreement, sensitivity and specificity for self-reported compared with operator-derived phone use by country and number of mobile phones

	Country	N	Weighted Kappa (95% CI)	Sensitivity (95% CI)		Specificity (95% CI)	
				High use[1]	Low use[1]	High use[1]	Low use[1]
Call frequency[1]							
One phone users	Finland	5820	0.30 (0.28-0.33)	38 (33-42)	81 (78-85)	97 (96-97)	76 (75-77)
	Sweden	25559	0.39 (0.38-0.40)	36 (34-37)	89 (88-90)	96 (95-96)	73 (73-74)
	UK	36568	0.33 (0.32-0.33)	40 (39-41)	85 (84-86)	95 (95-95)	71 (71-72)
	All	67947	0.35 (0.35-0.36)	38 (37-39)	87 (86-88)	95 (95-96)	72 (72-73)
Two (or more) phone users	Finland	377	0.30 (0.22-0.38)	47 (34-60)	80 (62-97)	93 (91-96)	76 (72-81)
	Sweden	5187	0.27 (0.25-0.29)	39 (37-42)	74 (68-79)	89 (88-90)	81 (80-82)
	UK	2317	0.33 (0.30-0.36)	50 (46-53)	67 (58-76)	91 (89-92)	79 (77-81)
	All	7881	0.30 (0.28-0.31)	43 (41-45)	72 (68-77)	90 (89-91)	80 (79-81)
Call duration[2]							
One phone users	Finland	5822	0.40 (0.38-0.42)	42 (39-44)	69 (66-73)	93 (92-93)	85 (84-86)
	Sweden	25582	0.53 (0.53-0.54)	54 (52-55)	81 (80-82)	92 (91-92)	84 (83-84)
	UK	36567	0.48 (0.47-0.48)	61 (60-62)	71 (70-72)	89 (89-89)	84 (83-84)
	All	67971	0.50 (0.49-0.50)	56 (55-56)	76 (75-76)	90 (90-91)	84 (84-84)
Two (or more) phone users	Finland	378	0.39 (0.33-0.45)	49 (40-58)	76 (58-94)	91 (88-95)	86 (82-90)
	Sweden	5191	0.40 (0.39-0.42)	57 (54-59)	67 (64-71)	85 (84-86)	84 (83-85)
	UK	2317	0.41 (0.39-0.44)	61 (58-65)	63 (58-68)	82 (80-84)	86 (84-87)
	All	7886	0.41 (0.39-0.42)	58 (56-59)	66 (64-69)	84 (84-85)	84 (84-85)

[1] Call frequency: High use ≥10 calls/day; Low use ≤6 calls/week. Call duration: High use ≥4 h/week; Low use <30 min/week.

Subgroup comparisons: sex and age group

Agreement between self-report and operator call frequency was significantly higher for men (κ=0.41, 95% CI: 0.40, 0.41) than women (κ=0.30, 95% CI: 0.29, 0.31), and increased across age strata (Table 4). Sensitivity of self-report for high call frequency was lower among women and young adults compared with men and older adults respectively (Table 4).

There was little difference in agreement (weighted kappa) between self-report and operator call duration according to sex or age strata (Table 4). For call duration, sensitivity of self-reported low call duration increased with increasing age (69% (95% CI: 68%, 70%), 77% (95% CI: 76%, 78%), 79% (95% CI: 78%, 80%) for 18-33 years, 34-49 years and ≥50 years, respectively), but the opposite was seen for high call duration as sensitivity decreased with increasing age (64% (95% CI: 63%, 66%), 58% (95% CI: 56%, 59%), 46% (95% CI: 44%, 47%) for 18-33 years, 34-49 years and ≥50 years, respectively). There were no sex differences in sensitivity for either low or high call duration

Table 4. Agreement, sensitivity and specificity for self-reported compared with operator-derived phone use, by age, sex, symptoms and concerns about mobile phone use (among one phone users only)

	Group	N	Weighted Kappa (95% CI)	Sensitivity (95% CI)		Specificity (95% CI)	
				High use[1]	Low use[1]	High use[1]	Low use[1]
Call frequency (3 categories)							
Sex	Men	29646	0.41 (0.40-0.41)	47 (46-49)	83 (82-85)	93 (93-94)	78 (77-78)
	Women	38233	0.30 (0.29-0.31)	28 (27-29)	90 (89-90)	97 (97-97)	68 (68-69)
Age	18-33yr	18075	0.29 (0.28-0.30)	30 (29-32)	90 (87-90)	97 (97-97)	68 (68-69)
	34-49yr	18923	0.35 (0.34-0.36)	41 (40-43)	84 (83-86)	95 (94-95)	74 (74-75)
	≥50yr	30805	0.39 (0.38-0.40)	41 (40-43)	87 (86-88)	95 (95-95)	74 (73-75)
Symptoms[2]	Yes	9714	0.34 (0.32-0.35)	42 (40-44)	81 (78-84)	94 (94-95)	78 (77-79)
	No	43487	0.36 (0.35-0.36)	38 (37-39)	88 (88-89)	96 (96-96)	70 (69-70)
Concern about mobile phone[3]	Yes	40295	0.34 (0.33-0.35)	38 (37-39)	86 (85-87)	95 (95-96)	74 (74-74)
	No	25439	0.37 (0.36-0.38)	39 (37-40)	88 (87-89)	96 (95-96)	70 (69-70)
Call duration (6 categories)							
Sex	Men	29661	0.49 (0.48-0.50)	55 (54-57)	75 (74-76)	90 (90-91)	84 (83-84)
	Women	38242	0.50 (0.50-0.51)	56 (54-57)	76 (75-77)	90 (90-91)	84 (84-85)
Age	18-33yr	18080	0.52 (0.51-0.53)	64 (63-66)	69 (68-70)	87 (86-87)	88 (87-88)
	34-49yr	18927	0.52 (0.51-0.53)	58 (56-59)	77 (76-78)	89 (89-90)	85 (85-86)
	≥50yr	30820	0.46 (0.46-0.47)	46 (44-47)	79 (78-80)	93 (92-93)	81 (80-81)
Symptoms[2]	Yes	9716	0.44 (0.43-0.46)	57 (55-59)	65 (62-67)	85 (85-86)	90 (89-90)
	No	43521	0.50 (0.49-0.50)	54 (53-55)	78 (77-79)	92 (92-92)	81 (80-81)
Concern about mobile phone[3]	Yes	40299	0.50 (0.49-0.50)	56 (55-57)	75 (74-76)	90 (90-90)	85 (85-85)
	No	25460	0.50 (0.49-0.51)	54 (52-55)	77 (76-78)	91 (91-92)	82 (81-83)

[1] Call frequency: High use ≥10 calls/day; Low use ≤6 calls/week. Call duration: High use ≥4 h/week; Low use <30 min/week. [2] Symptoms: 'Yes' defined as reported experiencing at least one (non-warming) health symptom in relation to mobile phone use. Those who reported warming sensations only were excluded from the analysis of symptoms. Total N included in symptoms analysis sums to less than the total number of one

mobile phone users due to excluding those who were missing data on symptoms (N=7495), reported warming sensation only (N=4939), reported contradictory answers (e.g. ticked the box “no symptoms” but then reported that they were experiencing certain symptoms when using a mobile phone) (N=2312), or were missing data on call frequency (N=140) or call duration (N=116), N.B. these Ns are not mutually exclusive.

[3] Concern about mobile phone use: ‘Yes’ defined as any level of concern regarding mobile phone use, and compared to those who expressed no concern about mobile phone use. Total N included in concerns analysis sums to less than the total number of one mobile phone users due to excluding those who were missing data on concerns (N=2213), or were missing data on call frequency (N=140) or call duration (N=116), N.B. these Ns are not mutually exclusive.

Subgroup comparisons: symptoms and risk perception

Agreement between self-reported and operator call duration was significantly lower among those who reported experiencing symptoms whilst (or shortly after) using a mobile phone (κ=0.44 (95% CI: 0.43, 0.46)) compared with those without symptoms (κ=0.50 (95% CI: 0.49, 0.50)), primarily because those with symptoms were more likely to overestimate low call duration (sensitivity=65% (95% CI: 62%, 67%) vs. 78% (95% CI: 77%, 79%) for those with and without symptoms respectively) (Table 4). A similar pattern was observed for call frequency, but the differences were smaller.

We observed little difference in the validity of either self-reported call frequency or call duration when comparing those concerned about the health effects of mobile phones vs. those without concerns, according to any of the measures (i.e. Kappa or sensitivity) (Table 4). Whilst there was a statistically significant difference in agreement between self-report and operator call frequency (κ=0.34, 95% CI: 0.33, 0.35 vs. κ=0.37, 95% CI: 0.36, 0.38 for concerned vs. no concern respectively), in absolute terms this difference is very small. Likewise there was no difference in the validity of either self-reported call frequency or call duration between those who reported concerns about either base stations or new technologies compared with those who did not (results not shown).

Sensitivity analyses

Results of subgroup analyses were similar, when repeated for two (or more) phone users, and when analyses excluded those who reported <10% of total use for the first phone, those who reported <40% of total use for the two phones, and those who reported regular use of their phone(s) by other people (results not shown).

Discussion

Main Findings

In this largest validation study to date, we found fair to moderate agreement between self-reported and operator-derived data on mobile phone use. The sensitivity of self-report was generally high for correctly identifying those with the smallest amount of mobile phone use, but lower for identifying heavy mobile phone use, in line with our observation that respondents in this study were more likely to underestimate than overestimate their mobile phone use. Subgroup analyses revealed that validity of self-reported mobile phone use differed according to sex, age, number of mobile phones and reported symptoms, but not according to risk perception regarding mobile phones. Users of two (or more) phones, and those who experienced symptoms during mobile phone use, were more likely to overestimate a small amount of mobile phone use compared with one phone users and those without symptoms.

Comparison with other studies

Compared to previous validation studies, our study found a similar proportion of respondents who misclassified their mobile phone use, in the order of 60%.(18, 38) Previous validation studies have demonstrated that subjects were prone to misclassify their mobile phone use by overestimating call duration ,(17, 18, 21, 38, 41-43) and suggest that the magnitude of overestimation, for both frequency and duration of calls, increases with increasing use.(38, 42) For example, a large published validation study (with 508 subjects from the Interphone case-control study), reported overestimation of mobile phone use by a factor of 4.64 among the heaviest users (>1640 hours of lifetime cumulative call time), but underestimation by a factor of 0.26 among lightest users (<5 hour of lifetime cumulative call time).(38) In contrast, our study suggests a tendency for underestimation of heavy mobile phone use (for both call duration and call frequency) within the COSMOS cohort. In the majority of previous validation studies, including those from the Interphone study,(18, 38) respondents reported their mobile phone use on a continuous scale,(17, 18, 21, 38) whereas the COSMOS questionnaire had categorical response options. It is possible that categorisation of mobile phone use can help to reduce overestimation in questionnaire data by truncating unrealistic or implausibly high usage estimates, a recurrent problem in previous studies of potential health effects of mobile phone use.(1, 17, 18, 38, 42, 43) Another explanation might be that in COSMOS participants were asked to report call duration per day or per week, whereas, in the Interphone (18, 38) and CEFALO (43) studies, for example, most or all participants reported call duration per call, and cumulative call duration was calculated as the product of call frequency and call duration per call. Hence, even if the call duration per call was only slightly overestimated, it could potentially lead to a considerable cumulative overestimation over the several months long validation study period. These data should also be interpreted in the context of temporal trends in mobile phone use, i.e. levels of mobile phone use, as measured in our study between 2007-2010, are likely to be higher compared to levels of mobile phone use when earlier validation studies were conducted.

Agreement between self-reported and operator call duration in this study was moderate (κ=0.50) but, nonetheless, considerably higher when compared to previous studies [e.g. κ=0.18,(21) 0.30,(37) and 0.45 (38)]. By virtue of access to operator data for many participants, COSMOS did not collect as detailed self-reported estimates of call duration as for example in the Interphone study. Therefore, it is likely that this observed improvement in validity compared to previous studies again arises from the use of categorical rather than continuous scale response options in self-reported call frequency and duration questionnaires.

Our findings demonstrate differential validity of self-reported mobile phone use according to sex, age, number of mobile phones, and self-reported experience of symptoms during mobile phone use. A few previous validation studies have alluded to population subgroup differences in validity of self-reported mobile phone use,(17, 21, 38, 42) but the evidence to date is inconsistent and based on very small numbers of participants. In contrast with our findings, a study of 68 adults reported better agreement between self-report and operator call frequency among women (κ=0.49) than men (κ=0.17).(21) Others have found no clear evidence for differences in validity of self-report exposure assessment by sex (38, 42), age (38, 42) or education.(38) One possible explanation for the demographic differences in validity observed in our study, could be differences in the use of mobile phones for work versus private/social purposes by age and by sex. This may influence the level of use, and also the accuracy of recall.

To our knowledge, this study is the first to investigate and quantify validity of self-reported mobile phone use among those who experience symptoms during mobile phone use or have concerns related to mobile phones.

Our findings demonstrate that those who experience symptoms when using a mobile phone are more likely to overestimate light mobile phone use, particularly call duration, compared to those without symptoms. This suggests that an individual's experience and/or perception of their health may influence the self-reporting of mobile phone use, likely affecting the validity of such exposure assessments. More specifically, it is possible that rumination bias (a form of information bias), whereby those with symptoms overestimate (consciously or subconsciously) their phone use in an effort to explain their symptoms, could be occurring in this subset of individuals. This finding has potential implications for the interpretation of previous cross-sectional studies investigating associations between mobile phone use and the symptoms reported here.(15, 16, 49) Overestimation of mobile phone use among those who report such symptoms would likely bias cross-sectional risk estimates away from the null, even if a true association does not exist,(30) thus potentially distorting any observed associations. We were unable to investigate whether the *severity* of symptoms influenced validity of self-reported mobile phone use in this study as we did not collect information on intensity, frequency or duration of symptoms. This should be explored in future research. We did not find a difference in validity when comparing those with no concern vs. any concern regarding mobile phone use.

Implications

In the past, many studies investigating the health effects of long-term mobile phone use have relied on self-report data to measure mobile phone use.(9, 14, 50) This is particularly true for the majority of case-control studies, where retrospective operator data was not available.(9, 14, 50)

However, self-report data continues to be valuable for newer studies in this field that adopt a prospective study design, such as the COSMOS cohort study. For these type of studies, it is possible to collect both self-report and operator data prior to the development of health outcomes, avoiding the potential problem of recall bias. Whilst operator data remain the gold standard measure of mobile phone use, these data have limitations nonetheless. Self-reported measures provide valuable information such as use of hands-free or lending/borrowing a mobile phone, that can supplement operator data, in order to facilitate a better understanding of an individual's mobile phone use. Furthermore, self-report data is particularly valuable when operator data is not available. This scenario is not uncommon in longitudinal studies, where an individual may change phone number, operator or country of residence, thereby precluding ongoing matching of operator data. In an international context, long-term operator data may not be available due to resource limitations or lack of willingness from network operators to provide these data for research purposes.

Our study demonstrates that there is considerable improvement to validity when a categorical, rather than continuous, scale is used to measure self-reported mobile phone. This highlights the ongoing value of using self-report data to measure mobile phone use.

Our findings also suggest that validity of self-report data, whilst still valuable to epidemiological research in this field, can be influenced by gender, age and the presence of symptoms. Therefore, it is important to understand the impact that demographic and health factors have on the validity of self-report data when interpreting subsequent epidemiological analyses.

Strengths and limitations

This is by far the largest study to date to investigate the validity of self-reported estimates of mobile phone use, and the first to report detailed subgroup comparisons, including those experiencing symptoms when using a mobile phone and/or concerns related to mobile phone use, in the general population. Our findings are likely to be more representative of the population at large than those of previous validation studies, which have largely been based on case-control studies of cancer risk.(21, 22, 38, 43) However, it is possible that the participants included in this validation study have a greater interest in the potential health effects of mobile phones and possibly, therefore, a greater awareness of their mobile phone use, than those who did not provide consent for their operator data to be accessed; this would likely result in underestimation of the true measurement error. In addition, mobile phone use over time is likely to be a highly dynamic phenomenon, dependant on a variety of technological and social factors. Therefore, current mobile phone use may differ from earlier periods in time when other validation studies were conducted.

The use of categorical response options for assessing mobile phone use in the COSMOS questionnaire can be considered both a strength and limitation of this study. Some information on inter-individual variation is lost through categorisation. However, as demonstrated in this study, the use of categorical response option may guide participants and prevent unrealistic responses and/or spurious precision, which may greatly misclassify mobile phone use, a recurrent problem in previous studies.(17, 18, 38, 43)

It is also important to recognise that operator data collected and reported by operators are primarily for the purposes of billing rather than scientific research. For example, this distinction becomes evident when considering how dropped calls are reported in the data; that is, calls that are disconnected due to signal loss or other technical issues, causing the caller to redial. The operator may record these as two separate calls and bill the user as such, whereas the caller may perceive this to be the one continuous call. However, in order to have an appreciable difference to our study findings, dropped calls would (a) need to occur often enough for study participants to misclassify their call frequency into another response category and (b) occur in different proportions between subgroups of study participants. In our opinion, these scenarios are unlikely and, therefore, dropped calls are unlikely to make any appreciable difference to our overall findings.

Operator data can also lead to exposure misclassification if linking or retrieved information is incorrect or if individuals regularly use a mobile phone, which is subscribed in someone else’s name or vice versa, if the phone used in operator linkage is used regularly by other people. We conducted several sensitivity analyses in an attempt to account for these potential sources of error and the results did not change, so any influence is likely to be small. Furthermore, the operator records may involve errors as they are extracted from different sources by the operators to incorporate all incoming and outgoing calls. For example, calls between two customers of the same network operator are sometimes counted only once, as some operators rely on billing records. Although efforts were made to obtain operator data for the three month period as close to the baseline as possible, there was some variation due to differences in operators' data storage protocols. Thus, in some instances, the three month period, for which operator data were obtained, was not always identical to the three month period for which self-reported data, or operator data on additional mobile phones, were obtained. It was assumed that mobile phone use would be relatively stable over these time intervals, but we cannot rule out the possibility that some disagreement in usage (particularly among those who used more than one mobile phone) could be attributed to these timing differences. Finally, direct comparison of the validity of self-reported call frequency compared with self-reported duration of calls is limited in this study, as the number of response categories differed for each variable (three for call frequency vs. six for call duration) and, thus, the level of agreement for each variable differs depending on the statistical method used (i.e. simple agreement indices favoured call frequency as the most accurate self-report parameter, whereas, for example, sensitivity for high use was greater for call duration compared to call frequency).(51)

Conclusions

Our findings support the ongoing use of self-report data in epidemiological research measuring mobile phone use. Furthermore, categorical response options in self-administered questionnaires appear to prevent unrealistically high self-reported mobile phone usage estimates. Whilst this may lead to some underestimation of heavy mobile phone use, the overall validity is greatly improved compared to questionnaires requiring participants to self-report call frequency and duration on a continuous scale. We recommend that self-reported mobile phone use is collected, but only prospectively, and not after disease has occurred. This study also demonstrates differences in validity of self-reported mobile phone use according to level of mobile phone use, and provides the first evidence for differences in validity of self-reported mobile phone use between those who do and do not experience symptoms while using a mobile phone. Studies investigating potential health effects of mobile phone use should consider taking these differential factors into account when interpreting risk estimates.

Acknowledgements

Above all, we thank all those participants who have joined the COSMOS cohort study. We thank mobile phone network operators in Denmark, Finland, the Netherlands, Sweden and the UK for allowing invitation of their subscribers and/or provision of operator traffic data. We thank Margaret Douglass, James Brook, Joe Gale, James Bennett, Samantha Udondem and Tom Kennett at Imperial College London for their contributions to the project. We thank Jonathan Pham at Imperial College London for helping to edit this paper. 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 is a National Institute for Health Research (NIHR) senior investigator and acknowledges support from the NIHR Biomedical Research Centre at Imperial College Healthcare NHS Trust and Imperial College London, and the NIHR Health Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-10141). 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.

Conflict of Interest

All authors declare they have no competing financial interests.

References

1. AGNIR. Health Effects from Radiofrequency Electromagnetic Fields. Report of the independent Advisory Group on Non-ionising Radiation. Health Protection Agency; 2012.

2. Ahlbom A, Feychting M, Green A, Kheifets L, Savitz DA, Swerdlow AJ. Epidemiologic evidence on mobile phones and tumor risk: a review. Epidemiology. 2009;20(5):639-52.

3. Pettersson D, Mathiesen T, Prochazka M, Bergenheim T, Florentzson R, Harder H, et al. Long-term mobile phone use and acoustic neuroma risk. Epidemiology. 2014;25(2):233-41.

4. Schoemaker MJ, Swerdlow AJ, Ahlbom A, Auvinen A, Blaasaas KG, Cardis E, et al. Mobile phone use and risk of acoustic neuroma: results of the Interphone case-control study in five North European countries. Br J Cancer. 2005;93(7):842-8.

5. Swerdlow AJ, Feychting M, Green AC, Leeka Kheifets LK, Savitz DA. Mobile phones, brain tumors, and the interphone study: where are we now? Environ Health Perspect. 2011;119(11):1534-8.

6. Lahkola A, Tokola K, Auvinen A. Meta-analysis of mobile phone use and intracranial tumors. Scand J Work Environ Health. 2006;32(3):171-7.

7. Frei P, Poulsen AH, Johansen C, Olsen JH, Steding-Jessen M, Schuz J. Use of mobile phones and risk of brain tumours: update of Danish cohort study. BMJ. 2011;343:d6387.

8. The INTERPHONE Study Group. Brain tumour risk in relation to mobile telephone use: results of the INTERPHONE international case-control study. International journal of epidemiology. 2010;39(3):675-94.

9. Coureau G, Bouvier G, Lebailly P, Fabbro-Peray P, Gruber A, Leffondre K, et al. Mobile phone use and brain tumours in the CERENAT case-control study. Occup Environ Med. 2014;71(7):514-22.

10. Hardell L, Carlberg M, Soderqvist F, Mild KH. Pooled analysis of case-control studies on acoustic neuroma diagnosed 1997-2003 and 2007-2009 and use of mobile and cordless phones. Int J Oncol. 2013;43(4):1036-44.

11. Hardell L, Carlberg M. Mobile phone and cordless phone use and the risk for glioma - Analysis of pooled case-control studies in Sweden, 1997-2003 and 2007-2009. Pathophysiology. 2015;22(1):1-13.

12. The INTERPHONE Study Group. Acoustic neuroma risk in relation to mobile telephone use: results of the INTERPHONE international case-control study. Cancer Epidemiol. 2011;35(5):453-64.

13. Takebayashi T, Akiba S, Kikuchi Y, Taki M, Wake K, Watanabe S, et al. Mobile phone use and acoustic neuroma risk in Japan. Occup Environ Med. 2006;63(12):802-7.

14. Hardell L, Carlberg M, Hansson Mild K. Pooled analysis of case-control studies on malignant brain tumours and the use of mobile and cordless phones including living and deceased subjects. Int J Oncol. 2011;38(5):1465-74.

15. Mortazavi SM, Ahmadi J, Shariati M. Prevalence of subjective poor health symptoms associated with exposure to electromagnetic fields among university students. Bioelectromagnetics. 2007;28(4):326-30.

16. Soderqvist F, Carlberg M, Hardell L. Use of wireless telephones and self-reported health symptoms: a population-based study among Swedish adolescents aged 15-19 years. Environ Health. 2008;7:18.

17. Parslow RC, Hepworth SJ, McKinney PA. Recall of past use of mobile phone handsets. Radiat Prot Dosimetry. 2003;106(3):233-40.

18. Vrijheid M, Cardis E, Armstrong BK, Auvinen A, Berg G, Blaasaas KG, et al. Validation of short term recall of mobile phone use for the Interphone study. Occup Environ Med. 2006;63(4):237-43.

19. Vrijheid M, Richardson L, Armstrong BK, Auvinen A, Berg G, Carroll M, et al. Quantifying the impact of selection bias caused by nonparticipation in a case-control study of mobile phone use. Ann Epidemiol. 2009;19(1):33-41.

20. Berg G, Schuz J, Samkange-Zeeb F, Blettner M. 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(3):217-24.

21. Samkange-Zeeb F, Berg G, Blettner M. Validation of self-reported cellular phone use. J Expo Anal Environ Epidemiol. 2004;14(3):245-8.

22. Pettersson D, Bottai M, Mathiesen T, Prochazka M, Feychting M. Validation of self-reported start year of mobile phone use in a Swedish case-control study on radiofrequency fields and acoustic neuroma risk. J Expo Sci Environ Epidemiol. 2015;25(1):72-9.

23. Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J. 2003;20(1):54-60.

24. Schulz KF, Grimes DA. Case-control studies: research in reverse. Lancet. 2002;359(9304):431-4.

25. Roosli M, Hug K. Wireless communication fields and non-specific symptoms of ill health: a literature review. Wien Med Wochenschr. 2011;161(9-10):240-50.

26. Frei P, Mohler E, Braun-Fahrlander C, Frohlich J, Neubauer G, Roosli M. Cohort study on the effects of everyday life radio frequency electromagnetic field exposure on non-specific symptoms and tinnitus. Environ Int. 2012;38(1):29-36.

27. Seitz H, Stinner D, Eikmann T, Herr C, Roosli M. Electromagnetic hypersensitivity (EHS) and subjective health complaints associated with electromagnetic fields of mobile phone communication--a literature review published between 2000 and 2004. Sci Total Environ. 2005;349(1-3):45-55.

28. Baliatsas C, Van Kamp I, Bolte J, Schipper M, Yzermans J, Lebret E. Non-specific physical symptoms and electromagnetic field exposure in the general population: can we get more specific? A systematic review. Environ Int. 2012;41:15-28.

29. Baliatsas C, Bolte J, Yzermans J, Kelfkens G, Hooiveld M, Lebret E, 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. 2015;218(3):331-44.

30. Armstrong BG. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occupational and environmental medicine. 1998;55(10):651-6.

31. Wacholder S, Hartge P, Lubin JH, Dosemeci M. Non-differential misclassification and bias towards the null: a clarification. Occup Environ Med. 1995;52(8):557-8.

32. Brenner H, Loomis D. Varied forms of bias due to nondifferential error in measuring exposure. Epidemiology. 1994;5(5):510-7.

33. Drews CD, Greeland S. The impact of differential recall on the results of case-control studies. International journal of epidemiology. 1990;19(4):1107-12.

34. Armstrong BG. The effects of measurement errors on relative risk regressions. Am J Epidemiol. 1990;132(6):1176-84.

35. White E. Design and interpretation of studies of differential exposure measurement error. Am J Epidemiol. 2003;157(5):380-7.

36. Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. International journal of epidemiology. 2005;34(3):680-7.

37. Schuz J, Johansen C. A comparison of self-reported cellular telephone use with subscriber data: agreement between the two methods and implications for risk estimation. Bioelectromagnetics. 2007;28(2):130-6.

38. Vrijheid M, Armstrong BK, Bedard D, Brown J, Deltour I, Iavarone I, et al. Recall bias in the assessment of exposure to mobile phones. J Expo Sci Environ Epidemiol. 2009;19(4):369-81.

39. Funch DP, Rothman KJ, Loughlin JE, Dreyer NA. Utility of telephone company records for epidemiologic studies of cellular telephones. Epidemiology. 1996;7(3):299-302.

40. Heinavaara 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. 2011;10:14.

41. Inyang I, Benke G, Morrissey J, McKenzie R, Abramson M. How well do adolescents recall use of mobile telephones? Results of a validation study. BMC Med Res Methodol. 2009;9:36.

42. Tokola K, Kurttio P, Salminen T, Auvinen A. Reducing overestimation in reported mobile phone use associated with epidemiological studies. Bioelectromagnetics. 2008;29(7):559-63.

43. Aydin D, Feychting M, Schuz J, Andersen TV, Poulsen AH, Prochazka M, et al. Predictors and overestimation of recalled mobile phone use among children and adolescents. Prog Biophys Mol Biol. 2011;107(3):356-61.

44. Schuz J, Elliott P, Auvinen A, Kromhout H, Poulsen AH, Johansen C, et al. An international prospective cohort study of mobile phone users and health (Cosmos): design considerations and enrolment. Cancer Epidemiol. 2011;35(1):37-43.

45. Toledano MB, Smith RB, Brook JP, 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. 2015;10(7):e0131521.

46. Toledano MB, Smith RB, Chang I, Douglass M, Elliott P. Cohort profile: UK COSMOS—a UK cohort for study of environment and health. International journal of epidemiology. 2015.

47. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968;70(4):213-20.

48. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

49. Sandstrom M, Wilen J, Oftedal G, Hansson Mild K. Mobile phone use and subjective symptoms. Comparison of symptoms experienced by users of analogue and digital mobile phones. Occup Med (Lond). 2001;51(1):25-35.

50. Group IS. Brain tumour risk in relation to mobile telephone use: results of the INTERPHONE international case-control study. International journal of epidemiology. 2010;39(3):675-94.

51. Brenner H, Kliebsch U. Dependence of weighted kappa coefficients on the number of categories. Epidemiology. 1996;7(2):199-202.