



Review-Bericht zur TNO-Studie; Version 1
03.02.2004 / Gregor Dürrenberger

Diskussionsstand zur TNO-Studie

1. Bedeutung der TNO-Studie

Die Bedeutung der Studie ist zweifellos die, dass erstmals Effekte von UMTS-Signalen auf den Menschen im Laborexperiment festgestellt wurden. In der Studie wurde vergleichend mit Fernfeldexpositionen (Typ Basisstation) sowohl von GSM 900 und GSM 1800 als auch von UMTS 2100 Signalen gearbeitet.

Es wurde eine schwache, aber statistisch signifikante Reduktion des Wohlbefindens beobachtet, sowohl in einer Gruppe mit Personen, die nach eigenen Angaben an unspezifischen Gesundheitsproblemen aufgrund von GSM-Basisstationen leiden, als auch in einer Referenzgruppe von Menschen ohne solche Leiden. Die Reduktion des Wohlbefindens betraf nur UMTS-Exposition, nicht GSM-Exposition.

Es wurden auch Auswirkungen auf die kognitiven Leistungen beobachtet, aber ohne klares Muster betreffen Expositionsart (GSM, UMTS), kognitive Teilfunktionen und Gruppenzugehörigkeit.

Weil die gefundenen Effekte bezüglich des Wohlbefindens vergleichsweise klein sind und ihre wirkliche Bedeutung für die Gesundheit unklar, müssen die Resultate durch unabhängige Replikations-Studien zuerst noch validiert werden. Die holländische Regierung hat erklärt, dass die Ergebnisse der TNO-Studie keine politischen oder regulatorischen Folgen haben, solange sie nicht wissenschaftlich verifiziert sind.

2. Download

http://www.tno.nl/en/news/article_6265.html

3. Studiendesign

Ansatz: Hypothesengeleitete Provokationsstudie, doppelblind, randomisiert, 3-fach cross-over

Exposition: (i) GSM 900 und (ii) GSM 1800; beide TDMA Modulation, Peak-Feldstärke der Signale am Standort der Versuchspersonen 1V/m, duty-cycle 50%, Kanalbelegung 4 in Folge, effektive Feldstärke am Standort der Versuchspersonen 0.7V/m, Fernfeldbedingungen. (iii) UMTS 2140, CDMA-Modulation, Feldstärke am Standort der Versuchspersonen 1V/m (peak und effektiv), 4 Frames, Fernfeldbedingungen. (iv) Sham-Exposition. Dauer der Expositionen: 30 Minuten in kontrollierter Umgebung (anechoic room).

Tests: 30 kognitive Tests am Computer zu Reaktionszeit, Erinnerungsvermögen, visuelle Aufmerksamkeit, dual-tasking, 2 Fragebogen zu Persönlichkeitsprofil (5 Variablen) und zum Wohlbefinden (3 Variablen und Summe). Dauer der kognitiven Tests 20 Minuten, Dauer des Fragebogens zum Wohlbefinden 10 Minuten. Dauer des Fragebogens zum Persönlichkeitsprofil (nur in der Trainings-Sitzung eingesetzt: 20 Minuten)

Versuchspersonen: 2 Gruppen (Elektrosensible, Referenzgruppe) à 36 Personen. Referenzgruppe (Durchschnittsalter 47, Frauenanteil 38%) nicht an die Gruppe der Elektrosensiblen angepasst (Durchschnittsalter 56, Frauenanteil 67%).

Statistische Aussagekraft: Die Samplegrösse für eine statistische Aussagekraft von 80%, basierend auf empirischen Daten zur interpersonellen und intrapersonellen Variabilität der kognitiven Tests bzw. zu Unterschieden im Summentotal beim Fragebogen zum Wohlbefinden, beträgt 28 Personen. Mit 36 Personen liegt das Sample über diesem Wert.

Sitzungs-Design: 1 Trainingssitzung mit allen Tests, dann randomisiert drei Expositionen: 2 real-Expositionen (aus dem Sample GSM 900, GSM 1800 und UMTS) und die Sham-Exposition. Diese Sitzungen mit den 4 kognitiven Tests und dem Fragebogen zum Wohlbefinden. Alle Sessions einer Person fanden innerhalb eines halben Tages (Morgen oder Nachmittag) statt. Dauer einer Sitzung 30 Minuten, wash-out Periode ca. 30 Minuten.



4. Resultate

Hypothese der Studie: es gibt keine statistisch signifikanten Unterschiede der kognitiven Leistungen und des subjektiven Wohlbefindens zwischen den 4 Expositionen.

Hauptresultat: Falsifikation der Hypothese.

- (i) Die Studie fand bei beiden Gruppen einen statistisch signifikanten Zusammenhang ($p=0.05$) zwischen UMTS-Exposition und Wohlbefinden. Bei der Gruppe der Elektrosensiblen waren die Resultate auf allen Variablen und beim Summentotal signifikant, bei der Referenzgruppe bei zwei Variablen und ebenfalls beim Summentotal. Die GSM-Expositionen zeigten gegenüber der Sham-Exposition keine statistisch signifikanten Unterschiede im Wohlbefinden.
- (ii) Bezüglich der kognitiven Leistungen waren die Ergebnisse heterogener. Signifikante Unterschiede ergaben sich bei den zwei Gruppen bei verschiedenen Tests und verschiedenen Expositionen. Zum Beispiel waren die Reaktionszeiten bei der Gruppe der Elektrosensiblen bei Exposition mit GSM 900 signifikant länger, bei der Referenzgruppe hingegen bei UMTS-Exposition. Beim Erinnerungsvermögen waren in der Gruppe der Elektrosensiblen die Testresultate nicht signifikant, bei der Referenzgruppe bei GSM 1800 und bei UMTS. Insgesamt zeigte die Referenzgruppe mehr signifikante Resultate.

Beurteilung durch die Autoren: Sie waren von den Ergebnissen überrascht, insbesondere von den Ergebnissen zum Wohlbefinden. Folgende Punkte geben sie für die Interpretation der Daten zu bedenken:

- (i) Vergleiche zwischen den 2 Gruppen können nicht gemacht werden. Nur gruppenspezifische Aussagen sind statistisch vertretbar.
- (ii) Der Fragebogen zum Wohlbefinden ist ein Auszug aus einem längeren Fragebogen, der in einer anderen Studie eingesetzt wurde. Die Ethikkommission hat alle für die vorliegende Untersuchung nicht relevanten Fragen streichen lassen. Es fehlte den Autoren die Zeit zur Validierung des gekürzten Fragebogens.
- (iii) Die Resultate zum Wohlbefinden sind zwar statistisch signifikant, aber nur schwach ausgeprägt.
- (iv) Aufgrund des Designs müssen zwei Limitationen bei der Interpretation berücksichtigt werden. Erstens kann die relativ kurze Expositionsdauer Effekte, die möglicherweise erst bei längerer Exposition entstehen, nicht erfassen. Zweitens ist es möglich, dass Effekte von einer Sitzung zur nächsten anhalten, da die Sitzungsintervalle verhältnismässig kurz sind.
- (v) Da es sich um eine erstmalige Studie handelt, ist eine Replikation durch ein anderes Labor unbedingt notwendig, um die Resultate zu verifizieren.
- (vi) Forschungsbedarf ist u.a. auf folgenden Gebieten angezeigt: Dosis-Effekt-Abhängigkeit, Bedeutung von Alter und Geschlecht, Effekte bei veränderten Signalcharakteristiken.

5. Kommentare und Kritik der Scientific Community, Antworten der Autoren

Statistik

Frage/Kritik: Aufgrund des Studiendesigns wurde jede Person zwei von drei möglichen Expositionen sowie der Sham-Exposition ausgesetzt. In der Auswertung werden die 36 Sham-Expositionen mit den jeweils 24 HF-Expositionen (2 GSM, 1 UMTS) verglichen, was statistisch nicht zulässig ist.

Antwort: Die Daten wurden aufgrund dieser Kritik korrekt analysiert. Bei einem kognitiven Test in der Referenzgruppe verschwand eine Signifikanz, alle anderen Befunde und Studienaussagen blieben dieselben (siehe Anhang 2).

Frage/Kritik: Die Gruppengrösse ist zu klein.

Antwort: Diese Kritik scheint unberechtigt. Die notwendige Gruppengrösse um mit 80%iger-Sicherheit einen signifikanten Unterschied ($p=0.05$) zwischen Sham-Exposition und realer Exposition zu erhalten, beträgt nach den Berechnungen der Autoren $n=28$. Die Versuche wurden mit einer Gruppengrösse von $n=36$



durchgeführt. Grössere Gruppen würden selbstverständlich die Aussagesicherheit (statistical power) erhöhen.

Frage/Kritik: Wenn die Mittelwerte, z.B. aus Tabelle 11.5, mit der Anzahl Probanden multipliziert werden, dann ergibt das Ergebnis keine natürliche Zahl, was es aber sein sollte.

Antwort: Das ist ein reines Rundungsproblem (siehe Anhang 1)

Frage/Kritik: Wie wurde die Signifikanz der einzelnen Fragen in Anhang E bestimmt?

Antwort: Die Berechnungen sind im Bericht nicht zu finden. Es wurden die Veränderungen der einzelnen Antworten einer Person unter realer Exposition gegenüber Sham-Exposition berechnet. Mit Hilfe eines linearen Modells wurden mögliche Effekten aufgrund unterschiedlicher Sitzungsreihenfolge korrigiert.

Frage/Kritik: Wie rechtfertigen sich die Annahmen, wann ein Unterschied als statistisch signifikant betrachtet wird?

Antwort: Bezüglich der kognitiven Tests wurden die Annahmen aus den Studien von Preece übernommen. Dort wird beispielsweise die Veränderung der Reaktionszeit um 5% als bedeutende Veränderung eingestuft. Bezüglich des Wohlbefindens wurden die Erfahrungen aus zwei klinischen Studien im Bereich Kardiologie übernommen, welche mit dem vollständigen Bulpitt-Fragebogen arbeiteten. Die Kritik wird akzeptiert, dass das was in einer klinischen Studie im Bereich Kardiologie relevant ist, nicht notwendigerweise auch in der vorliegenden Untersuchung relevant sein muss. Da trotz des gekürzten Fragebogens (23 statt 37 Fragen) derselbe Unterschied im Summentotal (2 Punkte) verwendet wurde, ist die Studie quantitativ betrachtet zumindest nicht restriktiver vorgegangen.

Interpretationshilfe: Konkret bedeuten 2 Punkte bei 23 Fragen mit Antworten in 4-stufiger Likert-Skala, dass zwei Fragen um eine Stufe schlechter oder eine Frage um zwei Stufen schlechter beantwortet wird als im Vergleichsfall (Sham-Exposition). Die Referenzgruppe hat im Summentotal 0.5 Punkte mehr bei UMTS-Exposition als bei Sham-Exposition. Das bedeutet, dass jede zweite Person im Schnitt eine Frage des ganzen Fragebogens um eine Stufe schlechter beantwortet hat. Bei der Gruppe der Elektrosensiblen beträgt die Differenz 3 Punkte. Jede Person hat 3 Fragen des Fragebogens um je eine Stufe schlechter beantwortet, oder eine Frage um 3 Stufen schlechter, etc.

Fragebogen

Frage/Kritik: Die Validität des Fragebogens zum Wohlbefinden scheint nicht geprüft worden zu sein.

Antwort: Diese Kritik ist berechtigt. Da der Fragebogen nicht validiert wurde, können keine Aussagen zur *Bedeutung* der angegebenen Veränderung im Wohlbefinden gemacht werden, sowohl innerhalb als auch zwischen den Gruppen.

Publikationsorgan

Frage/Kritik: Warum wurde die Studie nicht in einer wissenschaftlich begutachteten Fachzeitschrift eingereicht, um sich gegen Fehler und Schwächen vor einer Veröffentlichung abzusichern?

Antwort: Gegenwärtig wird ein Paper zur Einreichung an ein peer-reviewed journal vorbereitet. Die Ergebnisse wurden vor der Publikation von einem wissenschaftlichen Ausschuss und von Experten des United States Air Force Research Laboratory in Brooks City-Base, TX, kritisch durchgesehen. Die Autoren haben ihre Ergebnisse auf Verlangen der Regierung, welche die Studie finanzierte, publiziert. Die Studie wurde u.a. aufgrund eines parlamentarischen Vorstosses in Auftrag gegeben.



6. Vorschlag für eine Replikations-Studie

Hypothese

Ergebnisse der TNO-Studie werden durch Replikation falsifiziert.

Replikation

- Exposition: Beschränkung auf die Exposition mit statistisch signifikanten Resultaten im Bereich Wohlbefinden; also UMTS-Signal (zu prüfen: TNO- Signal oder IT'IS worst-case Signal?), identisches Expositions-set-up (anechoic room, Distanzen)
- Tests: Identischer Fragebogen zum Wohlbefinden (zu prüfen: Validierung; ev. andere oder zum Teil andere kognitive Tests)
- Sample: Ebenfalls 2 Populationen (Elektrosensible, Referenzgruppe; zu prüfen: Anpassung an die Gruppe der Elektrosensiblen)

Erweiterungen

- Exposition: zwei Levels. 1V/m (TNO) und 10V/m
- Tests: Zusätzliche Messungen von objektiven Parametern: EKG, Blutdruck, ev. EEG
- Sessions: eine Trainingssession, 2 Sham- und 2 Expositionssessions (1V/m und 10V/m). Das genaue Design muss noch geprüft werden. Ein oder zwei Tage? Randomisiert oder immer mit einer Sham-Session beginnend (wegen möglichen Langzeiteffekten)? Gegenüber der TNO-Studie ev. verlängerte Wartezeit zwischen den Sitzungen (längere wash-out Periode).
- Sample: Powerberechnung für minimale Samplegrösse um statistisch signifikante Resultate im erweiterten Design zu erhalten

Vorstudie

In zwei vom BAG finanzierten Vorstudien mit UNI-ZH (P. Achermann) und IT'IS (N. Kuster) die ca. März '04 abgeschlossen sein sollten, werden gegenwärtig die Designfragen bezüglich Exposition (Signalcharakteristik, levels), Tests (Auswahl kognitive Tests; EMV der zusätzliche Tests bzw. Messungen; Sessionplan), Sample (Grösse, matching criteria) geklärt sowie Kosten und Zeitplanung konkretisiert mit dem Ziel, ein wissenschaftlich robustes Proposal für eine Replikationsstudie vorlegen zu können. Austausch mit TNO/Holland besteht.

Kostenschätzung und Termine

In einer ersten groben Kostenschätzung wird von ca. 500KCHF ausgegangen, knapp die Hälfte der Kosten der TNO-Studie. Dabei value added: zusätzliche Messungen objektiver Parameter (EKG, Blutdruck, ev. EEG), sowie eines zweiten Expositions-niveaus. Die Studie dürfte ca. 1 Jahr dauern, frühester Beginn Frühsommer 2004.

Bedeutung

- (i) Die wissenschaftliche Bedeutung der vorgeschlagenen Replikations-Studie liegt in der Validierung der TNO-Resultate. Bis dato sind diese Befunde wissenschaftlich nicht gesichert (nicht reproduziert). Darüber hinaus wird die Studie wichtige zusätzliche Resultate liefern, insbesondere über mögliche Dosis-Wirkungs-Zusammenhänge und hinsichtlich der Wirkung der Expositionen auf messbare physiologische Parameter.
- (ii) Die gesellschaftliche Bedeutung der vorgeschlagenen Replikations-Studie bezieht sich auf die politische Diskussion und / bzw. auf die Regulierungsfrage. Zum einen begründen die Umweltverbände ihre Forderung nach einem UMTS-Moratorium mit den Ergebnissen der TNO-Studie. Zum zweiten sind die TNO-Ergebnisse für die Technologie-Bewertung durch die Bundesämter von Relevanz. Aber um ernsthaft konkrete Massnahmen zu diskutieren, falls solche angezeigt sein sollten, sind die Ergebnisse wissenschaftlich nicht ausreichend abgesichert. Dazu bedarf es einer qualitativ hochstehenden Replikations-Studie.

Finanzierung und Trägerschaft

Finanzierung: Es wird eine Mischfinanzierung angestrebt zwischen Bundesämtern, Industrie und weiteren Quellen.

Trägerschaft: Neutrale Trägerschaft (Auftraggeber der Studie) durch die Forschungsstiftung Mobilkommunikation, in Zusammenarbeit mit BAG und ev. BUWAL und/oder BAKOM.



Anhang 1

Original Table 11.5

Session	Group A				Group B				p-value* Group A vs Group B
	Mean	SEM	N	p-value** vs placebo	Mean	SEM	N	p-value** vs placebo	
Training	5,72	1,25	36		1,83	0,42	36		<0.05
Placebo	7,47	1,38	36		2,44	0,42	36		
900MHz	8,71	2,21	24	NS	2,25	0,45	24	NS	
1800MHz	7,33	1,19	24	NS	1,96	0,51	24	NS	
2100MHz	10,75	2,05	24	<0.05	3,08	0,70	24	<0.05	

Performing $5.72 \cdot 36 = 205,92$ gives rise to the suspicion that an error is made. This number should be an integer number. Analysis of this problem reveals that 5.72 represents 5,72222. Multiplying 5,72222 with 36 equals 205,99992 that obviously represent the integer value 206. From the next Table it is easily verified that this applies for the other numbers as well

Tabel 1. Modifications of Tabel 11.5 of TNO report FEL-03-C148 expressing more digits in the Mean values.

Session	Group A				Group B				p-value* Group A vs Group B
	Mean	SEM	N	p-value** vs placebo	Mean	SEM	N	p-value** vs placebo	
Training	5,72222	1,25	36		1,83333	0,42	36		<0.05
Placebo	7,47222	1,38	36		2,44444	0,42	36		
900MHz	8,70833	2,21	24	NS	2,2500	0,45	24	NS	
1800MHz	7,33333	1,19	24	NS	1,95833	0,51	24	NS	
2100MHz	10,7500	2,05	24	<0.05	3,08333	0,70	24	<0.05	

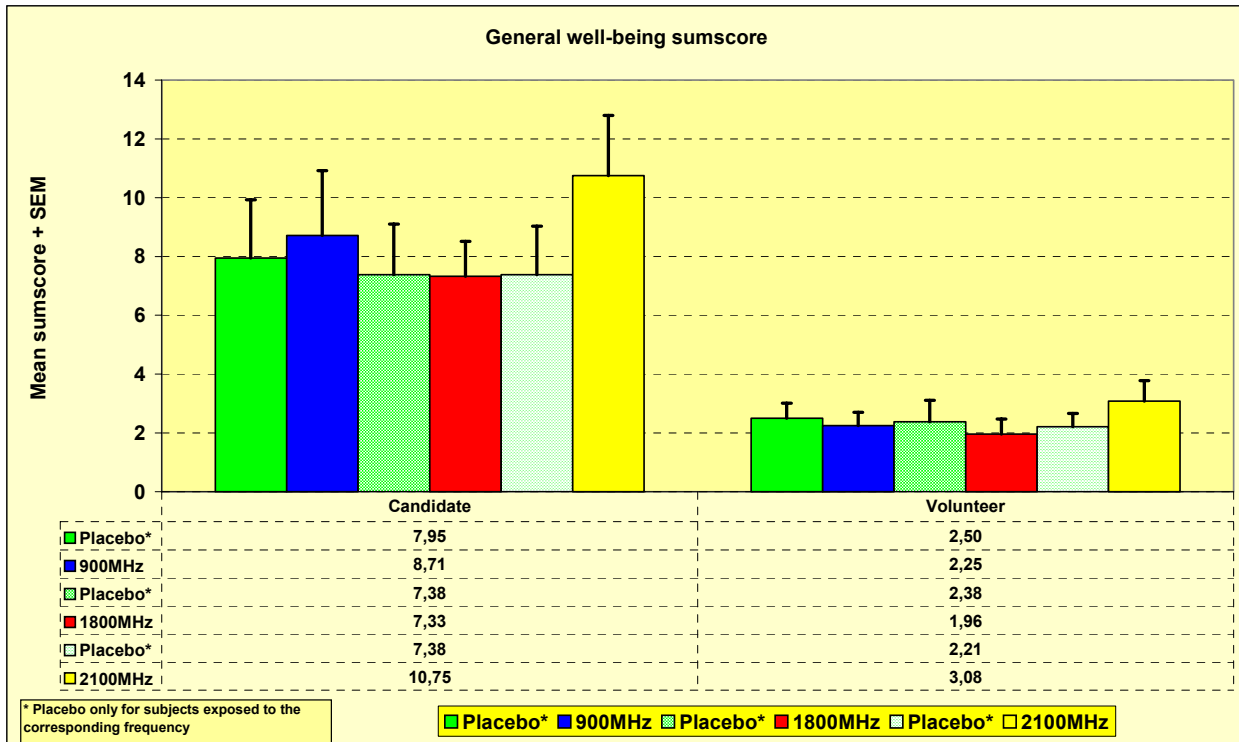
Anhang 2

2.1 General Well Being Sumscores

Tabel 2. General well-being sum score with all relevant decimals and comparison between corresponding frequency and pertaining placebo.

Session	Group A				Group B				p-value* Group A vs Group B
	Mean	SEM	N	p-value** vs placebo	Mean	SEM	N	p-value** vs placebo	
Training	5,72222	1,25	36		1,83333	0,42	36		<0.05
Placebo	7,47222	1,38	36		2,44444	0,42	36		
Placebo*	7.9580	1,98	24		2,5000	0,51	24		
900MHz	8,70833	2,21	24	NS	2,2500	0,45	24	NS	
Placebo*	7.37500	1,72	24		2,3750	0,73	24		
1800MHz	7,33333	1,19	24	NS	1,96583	0,51	24	NS	
Placebo*	7.37500	1,65	24		2,2080	0,45	24		
2100MHz	10,7500	2,05	24	<0.05	3,0833	0,70	24	<0.05	

* Placebo only for subjects exposed to the corresponding frequency



2.2 Cognitive Functions

Compared to the results presented in report FEL-03-C148, we lose one significance. This is the Memory Comparison Test, 1800 MHz within group B. Seven significancies remain and therefore the conclusions drawn for the cognitive functions remain.

Reaction Time test (intercept; msec)

Exposure	Group A				Group B				p-value* Group A vs. Group B
	Mean	SEM	N	p-value* vs. placebo	Mean	SEM	N	p-value* vs. placebo	
Placebo	1153	22,3	36		1139	23,2	36		NS
Placebo*	1136	22,9	24		1149	30,3	24		
900MHz	1196	34,6	24	P<0.05	1161	32,3	24	NS	
Placebo*	1172	30,9	24		1149	20,4	24		
1800MHz	1161	25,5	24	NS	1121	24,7	24	NS	
Placebo*	1152	27,7	24		1120	32,5	24		
2100MHz	1172	27,2	24	NS	1179	38,8	24	P<0.05	

* Placebo only for subjects exposed to the corresponding frequency



Memory Comparison test

Exposure	Group A				Group B				p-value* Group A vs. Group B
	Mean	SEM	N	p-value* vs. placebo	Mean	SEM	N	p-value* vs. placebo	
Placebo	27,8	2,9	35		26,4	2,8	36		NS
Placebo*	26.2	3.6	24		27.2	3.6	24		
900MHz	25,8	3,6	24	NS	23,3	2,6	24	NS	
Placebo*	26.1	3.1	23		24.4	2.6	24		
1800MHz	29,4	4,6	24	NS	20,2	4,2	24	NS	
Placebo*	31.5	3.5	23		27.5	3.9	24		
2100MHz	32,6	4,2	24	NS	20,7	3,8	24	P<0.05	

Visual selective attention Test

Exposure	Group A				Group B				p-value* Group A vs. Group B
	Mean	SEM	N	p-value* vs. placebo	Mean	SEM	N	p-value* vs. placebo	
Placebo	10,57	0,91	36		7,19	0,26	36		P<0.05
Placebo*	10,80	1,24	24		7,11	0,35	24		
900MHz	11,25	1,55	24	NS	7,01	0,35	24	NS	
Placebo*	10,71	1,04	24		7,26	0,33	24		
1800MHz	11,08	1,13	24	NS	7,29	0,34	24	NS	
Placebo*	10,20	1,07	24		7,21	0,26	24		
2100MHz	9,15	0,77	24	P<0.05	6,79	0,26	24	P<0.05	

Reaction parameter

Exposure	Group A				Group B				p-value* Group A vs Group B
	Mean	SEM	N	p-value* vs placebo	Mean	SEM	N	p-value* vs placebo	
Placebo	1304	37	36		1261	35	36		NS
Placebo*	1300	48	24		1282	46	24		
900MHz	1317	43	24	NS	1273	46	24	NS	
Placebo*	1317	49	24		1251	32	24		
1800MHz	1324	42	24	NS	1202	28	24	P<0.05	
Placebo*	1321	39	24		1250	47	24		
2100MHz	1336	41	24	NS	1247	35	24	NS	

Indicator for filtering irrelevant information

Exposure	Group A				Group B				p-value* Group A vs Group B
	Mean	SEM	N	p-value* vs placebo	Mean	SEM	N	p-value* vs placebo	
Placebo	192.0	37,5	34		124,7	30,5	36		NS
Placebo*	206,4	42,7	24		146.7	35.9	24		
900MHz	150,5	31,6	24	P<0.05	128,5	22,4	24	NS	
Placebo*	195,4	43,7	23		100,8	14,6	24		
1800MHz	220,5	37.6	24	NS	94,9	13.7	24	NS	
Placebo*	173,6	30,0	24		126,5	35,9	24		
2100MHz	168,9	20,9	24	NS	104,0	15,7	24	NS	

* Placebo only for subjects exposed to the corresponding frequency