Brain Topogr (2014) 27:731–746
DOI 10.1007/s10548-014-0355-9

ORIGINAL PAPER

A Reliability Study on Brain Activation During Active and Passive Arm Movements Supported by an MRI-Compatible Robot

Natalia Estévez · Ningbo Yu · Mike Brügger · Michael Villiger · Marie-Claude Hepp-Reymond · Robert Riener · Spyros Kollias

Received: 14 December 2013 / Accepted: 10 February 2014 / Published online: 10 April 2014

Abstract In neurorehabilitation, longitudinal assessment of arm movement related brain function in patients with motor disability is challenging due to variability in task performance. MRI-compatible robots monitor and control task performance, yielding more reliable evaluation of brain function over time. The main goals of the present study were first to define the brain network activated while performing active and passive elbow movements with an MRI-compatible arm robot (MaRIA) in healthy subjects, and second to test the reproducibility of this activation over time. For the fMRI analysis two models were compared. In model 1 movement onset and duration were included, whereas in model 2 force and range of motion were added to the analysis. Reliability of brain activation was tested with several statistical approaches applied on individual and group activation maps and on summary statistics. The activated network included mainly the primary motor cortex, primary and secondary somatosensory cortex, superior and inferior parietal cortex, medial and lateral premotor regions, and subcortical structures. Reliability analyses revealed robust activation for active movements with both fMRI models and all the statistical methods used. Imposed passive movements also elicited mainly robust brain activation for individual and group activation maps, and reliability was improved by including additional force and range of motion using model 2. These findings demonstrate that the use of robotic devices, such as MaRIA, can be useful to reliably assess arm movement related brain activation in longitudinal studies and may contribute in studies evaluating therapies and brain plasticity following injury in the nervous system.

Keywords fMRI · Elbow flexion/extension · Neurorehabilitation · MRI-Compatible robotic devices · Reliability · Sensorimotor network

Introduction

Functional magnetic resonance imaging (fMRI) allows measuring brain function in a non-invasive manner and

N. Estévez (✉) · S. Kollias
Department of Neuroradiology, University Hospital Zurich, Frauenklinikstrasse 10, CH-8091 Zurich, Switzerland
e-mail: estevez_natalia@hispeed.ch; Natalia.Estevez@usz.ch

N. Estévez
Institute of Social and Preventive Medicine, Zurich, Switzerland

N. Yu · M. Brügger · R. Riener
Sensory-Motor Systems Lab, Institute of Robotics and Intelligent Systems, ETH Zurich, TAN E 4 Tannenstrasse 1, CH-8092 Zurich, Switzerland

N. Yu
Institute of Robotics and Automatic Information System, Nankai University, Weijin Road 94, Tianjin 300071, China

M. Brügger
Institute for Biomedical Engineering, ETH Zurich and University of Zurich, Gloriastrasse 35, CH-8092 Zurich, Switzerland

M. Brügger
Center of Dental Medicine, University of Zurich, Zurich, Switzerland

M. Villiger · R. Riener
Spinal Cord Injury Center, University Hospital Balgrist, Forchstrasse 340, CH-8008 Zurich, Switzerland

M. Villiger · M.-C. Hepp-Reymond
Institute of Neuroinformatics, University of Zurich and ETH Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

therefore offers the possibility to repeat measurements over time. This is an important prerequisite to address questions related to brain reorganization after central or peripheral damage of the nervous system and to plasticity following training or rehabilitation treatments. In longitudinal studies, the use of paradigms able to provide robust activation across sessions is crucial. For example, during motor tasks differences in movement parameters across sessions (i.e. force, frequency, range of movement) may cause large differences in brain activation, complicating the interpretation of the results. To ensure a comparable motor performance across sessions, the relevant parameters of the task must be adequately controlled and monitored.

Consistency across sessions is even more challenging when studying patients with motor impairments whose motor output, i.e. force, range of movement etc., may change over time. This variability in task performance may consequently prevent meaningful conclusions related to brain activation changes following rehabilitative interventions and reorganization processes after injury.

MRI-compatible robotic devices have the potential to overcome the aforementioned limitations by providing control and monitoring of the motor performance over time. They guide the subjects to perform well-controlled and reproducible passive sensorimotor tasks and provide standardized conditions for active movement execution (Yu et al. 2008; for review see Tsekos et al. 2007). Furthermore, movement parameters can be recorded and quantified by the robotic system during the actual experiment. The collected data can then be incorporated into fMRI data analysis allowing accurate interpretations. Thus, MRI-compatible robots are promising tools for investigating brain reorganization mechanisms and plasticity related to neurorehabilitation by providing a well-controlled method for motor execution and for objectively monitoring the effect of therapy in patients with motor impairment.

For longitudinal assessments of brain function, test–retest analyses are essential to ensure that activation obtained with fMRI is reliable and does not randomly vary across repeated measures. In healthy subjects reliability of brain activation has been tested for a variety of cognitive and non-cognitive tasks (for review see Bennett and Miller 2010). With respect to motor function, reliability has been mainly assessed for active finger or hand movements (Carey et al. 2000; Friedman et al. 2008; Gountouna et al. 2010; Kimberley et al. 2008a, b; Kong et al. 2007; Lee et al. 2010; Loubinoux et al. 2001; McGregor et al. 2012; Yoo et al. 2007). In contrast, the reliability of brain activation patterns was rarely studied in passive motor tasks (Loubinoux et al. 2001). Only one study so far tested the reproducibility of activation in the primary motor cortex (M1) during active elbow flexion and extension (Alkadhi et al. 2002). Furthermore, to our knowledge there are no studies addressing reproducibility of passive arm movements. This is surprising, considering that arm movements are of major importance in the field of neurorehabilitation.

It is still a matter of debate which is the most appropriate test–retest analysis to assess reproducibility of brain activation. Therefore, different approaches were suggested, which all have advantages and disadvantages (for review see Bennett and Miller 2010). The calculation of various aspects of reliability should therefore give a more detailed estimation of the reproducibility in an fMRI study (Specht et al. 2003).

In the present investigation we test the reliability of brain activation during active and passive arm movements in healthy subjects. To this purpose an MRI-compatible arm robot (MaRIA), which guides extension and flexion of the elbow joint, was used in an fMRI event-related design (ERD) (Yu et al. 2008, 2011). The device allows monitoring and quantifying relevant movement parameters (movement onset, duration, force and range of motion). Here we present two possible fMRI models to show how this information can be best used to assess brain activation related to arm movements. This study had two main goals: first, to explore the brain network responsible for active and passive arm movements performed with MaRIA and second, to examine the reproducibility of this activation by applying various test–retest analyses. Since in future studies MaRIA will be used in various patient populations individual results are of major interest. Therefore, besides the reliability assessment on group results, the reproducibility of brain activation during active and passive arm movements was also tested at single subject level.

Materials and Methods

Participants

Nineteen healthy subjects (nine female, ten male, mean: 25 years, age range: 20–37 years) without history of neurological or psychiatric disorders were recruited for this study. All subjects had right-hand dominance (Annett 1970). The study was approved by the local ethics committee and all participants gave their written informed consent for participation prior to the experiment. In order to assess the reliability of arm movement related brain activation the volunteers participated in two fMRI sessions at intervals of 3–4 weeks.

MaRIA

MRI-compatible arm robot was developed by the Sensory Motor System Lab of the ETH Zurich (http://www.sms.

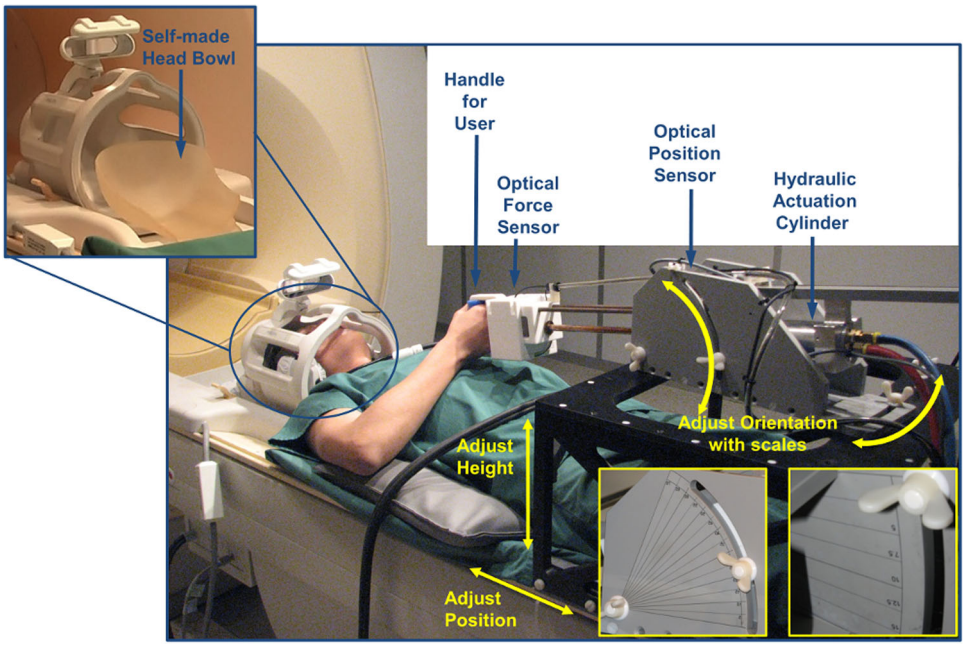

Fig. 1 Experimental setup: MaRIA is positioned slightly above the legs of the patients. At start position the arm is placed at 90° flexion. The position and orientation can be adjusted to fit the size of the patients. The settings used during the first session are stored and used in subsequent sessions. A self-made head bowl is used to avoid motion artifacts (modified from Yu et al. 2011 with permission of Springer Science and Business Media)

hest.ethz.ch/research/mr_robotics/setup). The device (Fig. 1) can be safely placed inside the MR scanner room, is compatible with fMRI, and allows extension and flexion movements of the elbow joint. A detailed description of this device was published in a pilot study (Yu et al. 2008, 2011). Therefore, only a brief description is provided here.

MRI-compatible arm robot allows adjustable, well-controlled, passive and active arm movements. It interacts with human subjects through a handle, which is attached to and driven by a hydraulic cylinder. The cylinder allows moving the handle in a translational direction, with a maximum motion range of 25 cm, maximum speed of 20 cm s and force up to 300 N. An optical force sensor, installed between the handle and the cylinder, measures the push and pull forces from the subject's arm to the cylinder. In addition an optical encoder measures the position of the handle, thus providing the recording of the handle's range of motion for each movement. The sensors also enable the assessment of movement onset and duration. This timing information allows an exact modeling of the brain activation related to arm movements. The position, height and orientation of the device constrain the movement of the robot and can be adjusted to fit the size of each subject. To further standardize the performance of the tasks the parameters used during one session are recorded for each subject and used in subsequent sessions. The device is controlled using MATLAB 7.6 (Mathworks Inc., Natick, MA, USA) and can be synchronized with other recording softwares, such as Presentation (http://www.neurobs.com/). Below we will refer to the range of motion of the device's handle as dROM.

fMRI Procedure and Experimental Paradigm

For the fMRI scans, the participants were positioned supine on the MR scanner table with the fixation frame of the device above the subjects' thighs. The participants were asked to flex the right elbow to reach the handle. The position, height and orientation of MaRIA were adjusted to ensure that subjects could reach the handle and perform the tasks in a comfortable way, while the upper arm remained close to the body without causing shoulder and head motion. The elbow was supported by a cushion for better comfort and stabilization of the upper arm. At the start position, the elbow was flexed by 90°. A maximal elbow extension reached approximately 120°, so that the range of motion of the subjects' elbow was about 30°.

To reduce head motion artifacts during data acquisition, we used a custom-made head support, which covered the top and partially the sides of the subjects' head (Hollnagel et al. 2011), thus limiting the range of head motion, especially in the cranio-caudal direction (Fig. 1). Additional foam pads restricted the motion in the left–right direction.

To investigate brain activation during the subjects' motor interactions with MaRIA, an ERD was used for the experiment. The experiment consisted of three conditions: passive arm movement, active arm movement and rest. In the passive condition, subjects were required to hold the device's handle and let it move without applying force. The speed was kept constant at 7.2 cm s. In the active condition, subjects had to push and pull the handle actively. The movement could only be initiated when the force reached a

certain threshold, defined as 20 % of the subjects' maximal voluntary push force (MVPF). The MVPF was measured by MaRIA for each subject in the scanner prior to fMRI scanning. Participants were instructed to push the fixed handle of the robot three times with their maximal voluntary force without moving head and body, and the mean force value was recorded. Above this threshold, an inverse viscous law was applied in such a way that an increase in the force applied by the subject induced an increase in the arm movement speed. Maximal speed was saturated to 10 cm s when the force reached 30 N or beyond. For both, active and passive movements, low speed and smooth movements were selected to avoid head motion and potential moving artifacts in the images (Yu et al. 2008, 2009). The dROM was approximately 16–20 cm depending on the body size of the individual subject. For each subject, the dROM and the linear movement trajectory remained the same for all passive and active movements. During the period of rest, subjects were simply asked to hold the device's handle without applying force. In order to test the reliability of this procedure in a standardized way, the same setting configuration used during the first session was applied in the second.

A total of 30 trials per condition were presented randomly to the participants. Each trial lasted 13.5 s and was composed of a short instruction followed by 8 s of task period and of an inter-stimulus interval with a jitter of 3 ± 1 s. The duration of the whole run was about 20 min. Passive and active movements were visually and acoustically guided to ensure that the active movements were performed similarly across trials and sessions, and had the same duration as the passive ones. Visual instructions, displayed on a screen in front of the subject, consisted of a green and a red square being presented for 4 s each, with the green always presented first. During the active condition, participants were instructed to push the device upon appearance of the green square and to pull it when the red one was displayed. The auditory instruction for the active condition consisted of the words "stossen" (German: "to push") and "ziehen" (German: "to pull"), which were synchronized with the green and red squares, respectively. During the rest and passive movement conditions the same colored squares were presented and the participants were asked to fixate the squares. For the passive and rest conditions the auditory instructions consisted of the words "stossen lassen" (German: "let it push") and "ziehen lassen" (German: "let it pull") and "Pause" (German: "pause"), respectively. The fMRI data acquisition and the tasks were synchronized applying Presentation (http://www.neurobs.com/). This software received trigger signals from the MR system and provided the visual and auditory instructions to the subjects. Additionally, it sent control commands to MaRIA instructing the device to switch from one condition to the other, allowing the initiation of active or passive movements. Prior to both scanning sessions the subjects were trained to practice the tasks outside of the scanner bore to ensure proper task performance.

During each scanning session the change in force and dROM, measured by the force and position sensors during the tasks, were displayed simultaneously in real time on a monitor outside the scanner room, allowing constant monitoring by the investigators to ensure that the subjects were performing the tasks correctly.

Behavioural Data Analysis

To assess the motor performance the following parameters were computed for each subject and session separately: force and dROM per trial, as well as mean force and mean dROM for the 30 active and passive movements separately.

During the arm movement itself, the force applied on the device's handle was normalized by the MVPF. In each session the mean force values were normalized by the respective MVPF.

The parameters for the individual trials were visually inspected to check whether the motor tasks were executed correctly. To identify differences between sessions, paired t-tests were performed on the normalized mean force for the active and the passive movements. The Kolmogorov–Smirnov test for mean dROM for active and passive movements showed significant results, indicating that the values were not normally distributed. Therefore, to test differences in the mean dROM between sessions, non-parametric tests were applied.

MRI Data Acquisition

The study was carried out in the MR-Center of the University and ETH Zurich, using a Philips Achieva 1.5 T MR system equipped with an eight channel SENSETM head coil. The functional acquisitions consisted of a T2* weighted, single-shot, field echo, EPI sequence of the whole brain ($TR = 3$ s, $TE = 50$ ms, flip angle $= 82°$, FOV $= 220 \times 220$ mm, acquisition matrix $= 128 \times 128$ mm, in-plane resolution $= 1.7 \times 1.7$ mm, slice thickness $= 4$ mm, SENSE factor 1.6). Additionally, anatomical images of the entire brain were acquired using a 3D, T1-weighted, field echo sequence ($TR = 20$ ms, $TE = 4.6$ ms, flip angle $= 20°$, in-plane resolution $= 0.9 \times 0.9$ mm, slice thickness $= 0.75$ mm, 210 slices).

Data Analysis

Image pre-processing and statistical analysis were performed using SPM8 (Welcome Department of Cognitive

Neurology, London, http://fil.ion.ucl.ac.uk/spm) implemented in MATLAB 7.6 (Mathworks Inc., Natick, MA, USA). "Realign and unwarp" facility was applied on the EPI images to correct for motion artifacts and additional susceptibility-by-movement interactions. The motion parameters obtained during this procedure were used to determine the extent of movements. Functional data that did not exceed displacement of one voxel size was included in the analysis. The realigned functional images of each session were then co-registered with the T1-weighted structural images acquired during the first MRI session. To achieve an accurate registration of the images between both scanning sessions DARTEL registration (Diffeomorphic Anatomical Registration using Exponentiated Lie algebra) was performed (Ashburner 2007). With this procedure the realigned EPI images were normalized and smoothed with an 8 mm full-with half-maximum Gaussian kernel. Additionally, a high-pass filter was applied on the preprocessed functional images to remove slow temporal drifts with a period longer than 128 s.

The statistical analysis was performed at single subject and group level. At the single subject level, the experimental conditions were modeled by the general linear model using two approaches: first by explicitly modeling all three conditions, i.e. rest, passive and active arm movements (contrasts against rest), and second by modeling only the movement conditions, i.e. active and passive arm movements (single contrasts). Additionally, for each of these approaches two different types of models were performed for each subject. In the first model, the experimental conditions were modeled in a more classical way using only information about the movement onset and duration. The exact movement onset and duration of each task, needed for modeling, were provided by the device and a canonical hemodynamic response function was used. In the second model, besides the three or the two experimental conditions respectively, two user defined regressors per session were added into the design matrix of each participant. The first one consisted of the mean applied force per scan normalized by the MVPF and the second was the maximal dROM per scan recorded by the device. This model should help to reduce additional variance due to differences in performance. All the analyses described below were performed for both models separately.

For both models individual statistical parametric maps (SPM) were calculated for each movement condition versus rest (first approach) and for the single contrasts for active and passive arm movements (second approach) for each session separately. Group analysis was performed according to the random effects analysis using the single subject contrast images as input. One-sample t-tests were performed for the four contrasts of interest per session. The significance level for the resulting statistical maps was set at $p < 0.05$, corrected for multiple comparisons (family wise error (FWE)). Additional analyses were performed at an uncorrected threshold of $p < 0.001$. Pair-t-tests were computed for the four contrasts to assess differences in activation maps across sessions,

Average and maximum t-values for each of the relevant contrasts were calculated in predefined anatomical regions of interest (ROIs) for both fMRI sessions separately. Differences in brain activation between the sessions were estimated by comparing the average t-value in each ROI using paired t-tests. The same analysis was also performed for the maximum t-value for each contrast and ROI. This analysis was performed in SPSS 19.0 (http://www.spss.com).

In the majority of the cases ROIs were defined based on probabilistic cytoarchitectonic maps implemented in the SPM anatomy toolbox (http://www.fz-juelich.de/ime/spm_anatomy_toolbox) (Eickhoff et al. 2005, 2006a, 2007). The bilateral analyzed areas were the M1, including Brodmann area (BA) 4a and b (Geyer et al. 1996), the primary somatosensory cortex (S1) including BA 3a, 3b, 1 and 2 (Geyer et al. 1999, 2000; Grefkes et al. 2001), and the secondary somatosensory cortex (S2) corresponding to the parietal operculum (OP1–4, Eickhoff et al. 2006a, b). Bilateral ROIs were also defined for the superior parietal cortex (SPC) including BA 5 and 7 (Scheperjans et al. 2008a, b) and inferior parietal cortex (IPC), comprising areas PFt, PF, PFm, PFcm, PFop, PGa, PGp (Caspers et al. 2006, 2008). The supplementary motor area (SMA) and the cingulate motor areas (CMA) were defined using the anatomic automatic labeling (Tzourio-Mazoyer et al. 2002) implemented in the standard software WFU Pickatlas (Maldjian et al. 2003). In order to define the premotor cortex (PMC) and divide it into a ventral and a dorsal part, a ROI for the BA 6 was created using the anatomy toolbox (Geyer 2004). Subsequently, the SMA was subtracted from the BA6 using MRIcron (http://www.mccauslandcenter.sc.edu/mricro/mricron/). The remaining part was divided into the dorsal PMC (PMd) and the ventral portion of BA6 which together with BA 44 was defined as the ventral PMC (PMv). Based on the meta-analysis by Mayka et al. (2006) the boundary between these two regions was set between z = 35 (MNI z = 38) medially and z = 45 (MNI z = 49) laterally. Finally, ROIs for the cerebellum (CB) were defined by combining all areas included in the anatomy toolbox (Diedrichsen et al. 2009).

Reliability Analyses

All reliability measures reported below were only performed in the ROIs that were activated in at least 80 % of the subjects, in all contrasts of interest and both sessions using both models. This allowed to reduce the data volume and to perform a reasonable comparison of the reliability

values across both models and conditions. These regions were the contralateral M1, S1, SMA, PMd, and SPC.

Reliability of Activation Maps

For comparison with other reliability studies, the relative amount of overlapping volume $R^{ij}_{overlap}$ between the two sessions was calculated according to the formula introduced by Rombouts et al. (1998):

$$R^{ij}_{overlap} = 2V_{overlap}/V_i + V_j \quad (1)$$

Where V_i and V_j denote the number of suprathreshold voxels within activation maps in session i and session j respectively, and $V_{overlap}$ represent the number of voxels that pass the threshold in both sessions. For the estimation of the $R^{ij}_{overlap}$ a statistical threshold of $p < 0.001$ (uncorrected for multiple comparisons) was used. The $R^{ij}_{overlap}$ can range from 0 (no overlap) to 1 (perfect overlap). This measure tests the reproducibility of the location of activated voxels above a threshold and is independent of the actual t-values of these voxels once they pass the threshold. In the present study, the $R^{ij}_{overlap}$ was used to assess test–retest reliability of brain activation of both the single subject data and the activation maps of the group analysis within predefine ROIs.

By setting a threshold, small differences in activation can be overestimated affecting considerably the size of the obtained $R^{ij}_{overlap}$. For example, some voxels may have a similar activation during both sessions, but may be below the threshold in one session and above it in the other. In spite of similar activation patterns these voxels would be classified as inconsistent between the sessions. To overcome this limitation, intraclass correlation coefficients (ICC) of contrast t-values for pairs of activation maps were calculated. This computation is based on all voxels in the brain and therefore, is not dependent on a threshold. In our study, test–retest reliability was computed across all voxels within each of the ROIs separately for individual and group activation maps. ICC values were calculated using a two-way mixed model ICC for consistency using the following formula (Shrout and Fleiss 1979):

$$ICC(3,1) = (BMS - EMS)/(BMS + (k-1)EMS) \quad (2)$$

BMS and EMS denotes the mean square for between voxel and error variance respectively, and k is the number of sessions. The ICC ranges from 0 (low reliability) to 1 (perfect reliability). Although some reliability studies have been performed on fMRI data in the past, there is still no consensus regarding the acceptable level of reliability. In order to have a basis for comparison in our study, ICC values were classified as 'excellent' above 0.75, 'good' between 0.59 and 0.75, 'fair' between 0.40 and 0.58 and 'poor' for values lower than 0.40, as proposed by Cicchetti and Sparrow (1981) In the following text 'high' will also be used for 'excellent' and 'moderate' for 'fair'. The calculated coefficient represents a value for intra-voxel reliability and we will refer to it as ICC_{within} (Raemaekers et al. 2007).

To summarize the results of the single subjects, the average $R_{overlap}$ and the average ICC_{within} were calculated. In order to average the ICC_{within} values across subjects, Fisher's z-transformation was applied on the ICC_{within} estimated for each subject.

Reliability of Summary Statistics

To assess test–retest reliability across subjects, ICC was also calculated on the average t-values and the maximum t-values for each ROI and contrast separately. ICC values were calculated using the same formula as before for the t-values of the individual and group activation maps (Shrout and Fleiss 1979). BMS and EMS denote the mean square for between subject and error variance respectively, and k denotes the number of sessions. In this case, the calculated coefficient represents a measure for between-subject reliability, referred as $ICC_{between}$. For this calculation, values are high for large between subject variance and small between session variance. The coefficients were tested against zero using a significance level of $p < 0.05$ (Shrout and Fleiss 1979).

Results

All 19 subjects accomplished the two fMRI sessions, but two (one female, one male) had to be excluded from the analysis, one due to the presence of significant movement artifacts and the other due to a technical problem in the synchronization of the tasks with the scanner.

Behavioral Performance

All subjects performed all active and passive movements as instructed. Mean MVPF was 47.2 N (±24.3) at the first and 42.4 N (±22.7) at the second session. The mean force for active movements was 20.0 N (±2.5) during the first and 17.8 N (±2.0) during the second session, while for passive movements the mean force was 3.5 N (±1.6) and 4.0 N (±1.4), respectively. Paired sample t-tests performed on the normalized force values for each movement condition and for MVPF did not show any significant differences in performance between sessions (passive, $t(16) = -1.29$, $p(16) = 0.21$; active, $t(16) = 0.33$, $p(16) = 0.75$; MVPF,

$t(16) = 2.1$, $p = 0.053$). In the active movement condition, the mean dROM was 17.1 cm (±1.8) during the first session and 18.3 cm (±2.1) during the second one. For passive movements, the mean dROM was 18.0 cm (±1) and 19.5 cm (±0.6), respectively. Furthermore, non-parametric tests on the dROM values did not differ significantly between sessions (passive, $z = -1.9$, $p = 0.61$; active, $z = -1.4$, $p = 0.15$).

Brain Activation

Model 1

In model 1, the experimental conditions were modeled using information about the movement onset and duration provided by the device.

In the first fMRI session, when contrasting the active movement condition with rest, group analysis revealed activation in left M1, S1, CMA, SPC, anterior insula and in the right anterior and posterior CB. Bilateral activation was found in S2, IPC, SMA, PMd, PMv and the mid insula ($p < 0.05$ corrected for multiple comparisons). During the second session, similar activation patterns were found, except in the PMv and in the left insula. Additionally, CMA was activated bilaterally. For both sessions, all reported areas were activated bilaterally when a less conservative correction was applied ($p < 0.001$ uncorrected for multiple comparisons). Additionally, activation was detected in the right middle temporal gyrus, bilaterally in the posterior insula and the basal ganglia, and in the left thalamus and brainstem (Fig. 2a). For the single contrast, active movement activation was found left in M1, S1, SMA, PMd, SPC, bilaterally in S2, IPC, and in the right PMv, CMA and anterior CB during the first session. During the second session this first model showed activation only in left M1, S1, SMA, PMd, SPC and right in IPC ($p < 0.05$ corrected for multiple comparisons). In both sessions, non-corrected activation maps revealed activation in the same network as for the active movement condition contrasted with rest with the exception of the left thalamus, right basal ganglia and right middle temporal gyrus (Fig. 2b).

When contrasting passive movement with rest for both fMRI sessions, the group activation patterns were similar to those in the contrast active movement versus rest. Only the insula and the PMv were not activated. In addition, activation was found in the left anterior CB during the first session. PMd was activated during the first session bilaterally and only on the left during the second one. Bilateral activation was found in CMA during both sessions ($p < 0.05$ corrected for multiple comparisons). For both sessions, all areas of this network showed bilateral activation when the activation maps were not corrected for multiple comparisons ($p < 0.001$). Additional activation was detected in the left thalamus and in the basal ganglia, middle temporal gyrus, PMv and the mid and posterior insula bilaterally (Fig. 2e). For the single contrast, passive movement activation was found in the left M1, S1, PMd and IPC in both sessions. Activation in S2 was only detected in the left hemisphere during the first session. When activation maps were not corrected for multiple comparisons ($p < 0.001$) the same activation pattern was found as for the contrast of passive movements versus rest, except for the right M1, S1, SPC, PMv and left CB (Fig. 2f).

For both active and passive movements, the single contrast showed in general less activation when compared to the contrast with rest. Coordinates for local maxima for all contrasts and ROIs using model 1 are shown in Table 1.

Model 2

In this model, besides the experimental conditions, additional movement parameters (i.e., force and dROM) provided by the device were implemented into the data analysis.

Applying model 2, the active movement condition compared to rest showed for both sessions the same activation patterns as in the analysis with the first model. This was the case using both thresholds ($p < 0.05$ corrected and $p < 0.001$ uncorrected for multiple comparisons, Fig. 2c). For both fMRI sessions the single contrast for active movements revealed activation in left M1, S1, SMA, CMA, PMd, SPC, in S2, IPC bilaterally, and in right PMv, and right posterior CB. During the second session activation was also found in the right mid insula and CMA ($p < 0.05$ corrected for multiple comparisons). Uncorrected activation maps revealed for both sessions the same network as in the active movement condition contrasted with rest, except for the right middle temporal gyrus (Fig. 2d).

For both sessions and thresholds the activation patterns in the passive movement condition compared to rest activation were similar to those reported for model 1 ($p < 0.05$ corrected and $p < 0.001$ uncorrected for multiple comparisons, Fig. 2g). For the single contrast passive movement activation was found in the same network as in the contrast with rest, except for the bilateral activation in SMA and CMA during the first session. Using this second model, the same activation patterns as those for passive movement condition contrasted with rest were found when the activation maps were not corrected for multiple comparisons ($p < 0.001$, Fig. 2h).

For active and passive movement, the activation pattern of the contrast with rest and the single contrast were largely

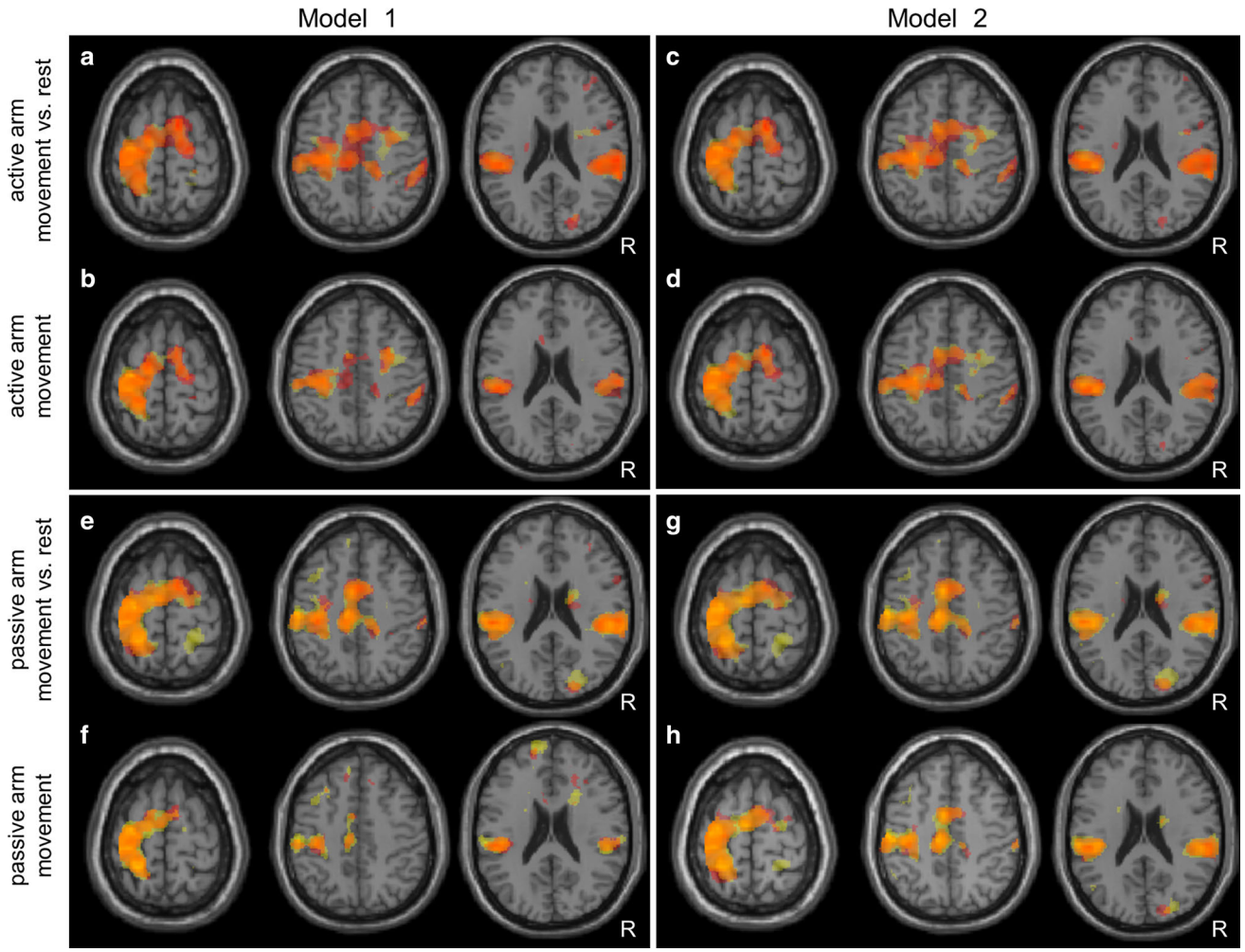

Fig. 2 Transversal sections showing the overlap of activation in both fMRI sessions for all contrasts of interest and for model 1 (*a*, *b*, *e*, *f*) and model 2 (*c*, *d*, *g*, *h*) ($p < 0.001$ uncorrected for multiple comparisons). Activation during first session (*red*), second session (*yellow*) and in both sessions (*orange*) were superimposed on a single subject template using xjView (http://people.hnl.bcm.tmc.edu/cuixu/xjView/). The most informative slices are displayed

identical. Coordinates for local maxima for all contrasts and ROIs using model 2 are shown in Table 2.

Systematic Changes in Brain Activation

For both models and all contrasts of interest, paired-t-tests analysis computed on the activation maps did not reveal any significant differences between sessions ($p < 0.05$ corrected for multiple comparisons). Additionally, no significant differences were found on average t-values for all the contrasts in the predefined ROIs. For the ROI analyses significant differences were only found on maximum t-values for the single contrast of active movements in contralateral M1 and S1 using model 1 ($p < 0.05$ non-corrected for multiple comparisons). For all other contrasts of interest and for model 2 no significant differences were found on maximum t-values.

Reliability Analyses

Reliability of Activation Maps

Overlap Ratios ($R_{overlap}$) The averages $R_{overlap}$ of the single subjects are presented in Table 3 for the two models. For both models the contrasts of the movement conditions with rest showed good to excellent reliability for activation in M1, S1, and PMd. Reliability ranged from moderate to good for SMA and moderate for SPC. In all ROIs except for the SPC, the $R_{overlap}$ calculation revealed slightly higher values for the active movement condition compared to rest using model 1 than with model 2. The opposite was observed for the passive movement condition against rest. For both single contrasts (i.e. active and passive arm movements), reliability was mainly good when modeling the data with model 1, only the SMA and SPC showed

Table 1 Coordinates of local maxima (MNI) for all ROIs and contrasts of interest during the first and second session using model 1

Model 1													
Active arm movement													
ROI		Contrast with rest						Single contrast					
		Session 1			Session 2			Session 1			Session 2		
		x	y	z	x	y	z	x	y	z	x	y	z
M1	L	**−32**	**−27**	**60**	**−33**	**−21**	**57**	**−27**	**−21**	**53**	**−33**	**−21**	**59**
	R	12	−30	50	9	−29	48	20	−26	57	26	−33	65
S1	L	**−33**	**−30**	**59**	**−30**	**−32**	**59**	**−33**	**−30**	**59**	**−32**	**−30**	**62**
	R	17	−35	50	**36**	**−27**	**38**	20	−39	56	36	−27	38
SMA	L	**−8**	**−6**	**54**	**−14**	**−12**	**65**	**−12**	**−11**	**53**	**−15**	**−11**	**63**
	R	**12**	**1**	**66**	**12**	**0**	**65**	15	−11	66	14	0	62
CMA	L	**−6**	**3**	**42**	**−8**	**1**	**44**	−8	−6	50	−8	1	41
	R	17	−30	42	**11**	**−29**	**44**	**18**	**−30**	**42**	**15**	**−29**	**41**
PMd	L	**−27**	**−21**	**60**	**−30**	**−18**	**57**	**−29**	**−20**	**56**	**−29**	**−20**	**56**
	R	**21**	**−17**	**65**	**35**	**−3**	**45**	20	−18	65	39	−3	44
PMv	L	**−44**	**9**	**6**	−42	−6	50	−50	1	6	−48	3	6
	R	**54**	**7**	**9**	48	9	8	**54**	**6**	**8**	44	−3	44
SPC	L	**−21**	**−41**	**62**	**−18**	**−42**	**63**	**−18**	**−39**	**63**	**−20**	**−41**	**65**
	R	15	−29	41	11	−29	44	17	−29	42	15	−29	41
IPC	L	**−51**	**−30**	**23**	**−51**	**−30**	**23**	**−51**	**−30**	**23**	−51	−29	23
	R	**60**	**−26**	**23**	**57**	**−26**	**30**	**57**	**−32**	**41**	**63**	**−27**	**35**
S2	L	**−48**	**−30**	**23**	**−44**	**−32**	**23**	**−48**	**−30**	**23**	−50	−29	23
	R	**62**	**−26**	**23**	**56**	**−27**	**26**	**62**	**−24**	**23**	44	−29	26
CB	L	2	−65	−16	−2	−48	−24	0	−51	−26	0	−50	−24
	R	**20**	**−54**	**−20**	**21**	**−50**	**−23**	**9**	**−53**	**−15**	21	−53	−21
Passive arm movement													
ROI		Contrast with rest						Single contrast					
		Session 1			Session 2			Session 1			Session 2		
		x	y	z	x	y	z	x	y	z	x	y	z
M1	L	**−32**	**−26**	**59**	**−33**	**−32**	**56**	**−33**	**−26**	**57**	**−33**	**−27**	**66**
	R	2	−21	50	0	−26	50						
S1	L	**−33**	**−30**	**59**	**−32**	**−33**	**59**	**−33**	**−30**	**59**	**−24**	**−41**	**57**
	R	20	−33	47	24	−44	65				23	−41	57
SMA	L	**0**	**3**	**47**	**−8**	**−6**	**56**	−12	−6	71	−8	−11	74
	R	**11**	**3**	**68**	**2**	**−3**	**53**	11	0	69	6	−5	59
CMA	L	**−8**	**−23**	**47**	**−12**	**−26**	**41**	−6	−18	48	−9	−21	44
	R	**3**	**3**	**44**	**12**	**7**	**38**	12	27	18	14	9	38
PMd	L	**−35**	**−27**	**69**	**−35**	**−27**	**69**	**−35**	**−27**	**69**	**−35**	**−27**	**69**
	R	**3**	**3**	**44**	**0**	**−24**	**47**	0	−17	53	0	−17	53
PMv	L	−50	1	6	−44	−8	53	−50	1	6	−44	−12	53
	R	57	7	8	53	3	0						
SPC	L	**−23**	**−44**	**62**	**−18**	**−42**	**63**	−23	−42	62	**−24**	**−42**	**66**
	R	17	−35	44	14	−27	45				**20**	**−53**	**60**
IPC	L	**−51**	**−29**	**21**	**−51**	**−32**	**20**	**−51**	**−29**	**23**	**−59**	**−29**	**26**
	R	**60**	**−33**	**23**	**60**	**−35**	**24**	54	−27	29	53	−32	24
S2	L	**−50**	**−30**	**20**	**−44**	**−27**	**20**	**−45**	**−30**	**21**	−44	−26	21
	R	**60**	**−26**	**24**	**53**	**−29**	**24**	56	−27	26	53	−29	24

Table 1 continued

Passive arm movement													
ROI		Contrast with rest						Single contrast					
		Session 1			Session 2			Session 1			Session 2		
		x	y	z	x	y	z	x	y	z	x	y	z
CB	L	**−33**	**−48**	**−33**	2	−60	−14	0	−69	−6	−14	−62	−9
	R	**26**	**−50**	**−21**	**21**	**−48**	**−21**	17	−56	−12	24	−53	−20

Bold denotes activations corrected for multiple comparisons with FWE $p < 0.05$; non-bold denotes uncorrected activations with a threshold of $p > 0.001$

M1 primary motor cortex, *S1* primary somatosensory cortex, *SMA* supplementary motor area, *CMA* cingulate motor areas, *PMd* dorsal premotor cortex, *PMv* ventral premotor cortex, *SPC* superior parietal cortex, *IPC* inferior parietal cortex, *S2* secondary somatosensory cortex, *CB* cerebellum

moderate values. For model 2, the $R_{overlap}$ values were higher for both conditions in all ROIs than using model 1. This was especially the case for the passive movement condition.

For group activation, all ROIs showed high reliability using both models (see Table 3). Analog to the single subjects' data, group activation maps showed mainly higher reliability for both single contrasts when controlling for motor performance.

Intra Class Correlation (ICC_{within}) Average ICC values for single subject and ICC values obtained for group activation maps are given in Table 3. For single subjects the intraclass correlation of t-values between the two sessions showed high reliability in M1, S1, and PMd and good reliability in SMA and SPC for all contrasts of interest and both models. Analog to the calculation of $R_{overlap}$, model 2 yielded better reliability for the single contrast of passive movements. For all contrasts and using both models group results were found to be highly reproducible for all ROIs.

Reliability of Summary Statistics

Intra Class Correlation ($ICC_{between}$)

Results for the ICC on average and maximum t-values are presented in Table 4. For the active movement condition in both, contrasts with rest and single contrasts, good to excellent reliability was found. The ICC values were significant in all ROIs analyzed with both models. ICC values were mainly higher for model 1 than for model 2.

The contrasts using passive movements showed low to good reproducibility. For model 1, the passive condition compared to rest showed significant values for M1, S1 and SMA, but not for PMd and SPC for average t-values. For the single contrast, intraclass correlations were only significant in M1 and SMA. However, using model 2, average t-values for all ROIs, except PMd, showed moderate but significant ICC values for both contrasts of passive movements (i.e. single contrast and contrast with rest), suggesting that this model improves the reliability of activations. For maximum t-values, all ROIs showed significant intraclass correlations in both models. Only intraclass correlation of SMA was not significant for both models and SPC for the first one.

Discussion

This study explores the brain network activated by active and passive elbow movements performed with the support and guidance of an MaRIA and tests the reproducibility of this activation. Brain activation was found in expected areas of the sensorimotor network for elbow movements and was reliable across sessions at single subject and group level. Thus, this device may allow longitudinal assessments of brain function in healthy subjects and potentially, in future studies on patients.

This outcome was possible assessing the following methodological approach. Quantitative data of the movement performance—onset, duration, force and dROM—provided by the robot were used to analyze the fMRI data. Two models were tested. With the first (model 1), the movement onset and duration were incorporated into the data analysis, allowing precise modeling of the performed movement. In the second approach (model 2), force and dROM were additionally implemented in the analysis as regressors removing variance in movement performance between trials. In order to provide a detailed estimation of the reproducibility of brain activation acquired with these approaches several statistical methods were applied on individual and group data.

For active movements, both models exhibited brain activation in a network including mainly the primary sensorimotor cortex (M1 and S1), secondary somatosensory cortex, insula, superior and inferior parietal lobules and medial and lateral premotor areas. Additionally, activation was found in anterior and posterior cerebellum, basal ganglia, thalamus and brain stem. These findings are largely consistent with earlier investigations of simple elbow

Table 2 Coordinates of MNI for all ROIs and contrasts of interest during the first and second session using model 2

Model 2													
Active arm movement													
ROI		Contrast with rest						Single contrast					
		Session 1			Session 2			Session 1			Session 2		
		x	y	z	x	y	z	x	y	z	x	y	z
M1	L	**−30**	**−20**	**53**	**−32**	**−20**	**54**	**−30**	**−20**	**53**	**−32**	**−20**	**54**
	R	12	−30	51	11	−29	48	12	−30	51	11	−29	48
S1	L	**−33**	**−30**	**59**	**−30**	**−32**	**60**	**−33**	**−30**	**59**	**−32**	**−30**	**62**
	R	17	−35	50	33	−29	39	17	−33	50	36	−27	38
SMA	L	**−8**	**−6**	**56**	**−14**	**−12**	**63**	**−8**	**−8**	**56**	**−14**	**−12**	**63**
	R	**12**	**0**	**66**	**12**	**−2**	**66**	**12**	**0**	**65**	**12**	**−2**	**66**
CMA	L	**−8**	**3**	**42**	**−8**	**3**	**44**	**−8**	**3**	**42**	**−8**	**3**	**44**
	R	17	−30	44	**11**	**−29**	**45**	17	−30	44	**12**	**−27**	**45**
PMd	L	**−26**	**−20**	**59**	**−32**	**−18**	**60**	**−26**	**−20**	**62**	**−33**	**−18**	**59**
	R	**20**	**−18**	**65**	17	−12	62	20	−18	65	14	−8	63
PMv	L	**−45**	**9**	**3**	−48	3	6	−50	1	6	−44	−12	53
	R	**56**	**7**	**11**	50	7	6	**56**	**7**	**9**	50	7	6
SPC	L	**−14**	**−26**	**48**	**−18**	**−42**	**63**	**−14**	**−26**	**48**	**−18**	**−42**	**63**
	R	17	−30	44	11	−29	45	17	−30	44	11	−29	47
IPC	L	**−51**	**−30**	**23**	**−42**	**−32**	**21**	**−50**	**−32**	**23**	**−42**	**−32**	**21**
	R	**51**	**−27**	**32**	**56**	**−27**	**30**	**51**	**−27**	**32**	**51**	**−26**	**29**
S2	L	**−48**	**−30**	**23**	**−42**	**−32**	**23**	**−48**	**−30**	**23**	**−44**	**−30**	**21**
	R	**62**	**−26**	**24**	**44**	**−24**	**26**	**62**	**−24**	**23**	**44**	**−24**	**26**
CB	L	0	−71	−7	2	−63	−14	0	−62	−18	2	−63	−14
	R	**25**	**−48**	**−25**	**21**	**−50**	**−23**	**25**	**−48**	**−25**	**21**	**−50**	**−23**
Passive arm movement													
ROI		Contrast with rest						Single contrast					
		Session 1			Session 2			Session 1			Session 2		
		x	y	z	x	y	z	x	y	z	x	y	z
M1	L	**−32**	**−27**	**60**	**−33**	**−32**	**56**	**−33**	**−29**	**59**	**−32**	**−29**	**62**
	R	2	−21	50	0	−26	50	2	−21	50	2	−23	48
S1	L	**−33**	**−30**	**59**	**−32**	**−32**	**59**	**−33**	**−30**	**59**	**−33**	**−32**	**59**
	R	33	−38	53	24	−42	66	33	−35	56	32	−38	53
SMA	L	**0**	**3**	**47**	**−8**	**−20**	**50**	−3	−3	56	**−8**	**−20**	**50**
	R	**11**	**0**	**66**	**2**	**−3**	**53**	11	0	66	**6**	**−2**	**60**
CMA	L	**−2**	**0**	**47**	**−3**	**0**	**47**	−8	−23	48	**−8**	**−23**	**45**
	R	**3**	**3**	**44**	**12**	**7**	**38**	2	1	44	11	7	39
PMd	L	**−33**	**−26**	**71**	**−35**	**−27**	**69**	**−33**	**−26**	**71**	**−35**	**−27**	**69**
	R	**0**	**0**	**47**	0	0	47	0	0	47	0	0	47
PMv	L	−54	7	14	−44	−8	53	−50	1	6	−44	−9	53
	R	63	11	5	62	11	5	63	11	5	62	14	3
SPC	L	**−23**	**−50**	**71**	**−18**	**−42**	**63**	−21	−50	71	**−24**	**−44**	**68**
	R	17	−35	44	14	−27	45	17	−33	42	21	−44	68
IPC	L	**−51**	**−29**	**21**	**−51**	**−29**	**20**	**−51**	**−30**	**23**	**−59**	**−26**	**21**
	R	**60**	**−35**	**23**	**60**	**−35**	**24**	**60**	**−29**	**24**	**53**	**−29**	**23**
S2	L	**−50**	**−30**	**20**	**−42**	**−29**	**18**	**−47**	**−30**	**21**	**−44**	**−29**	**20**
	R	**45**	**−30**	**21**	**53**	**−29**	**24**	**56**	**−27**	**26**	**53**	**−29**	**24**

Table 2 continued

Passive arm movement													
ROI		Contrast with rest						Single contrast					
		Session 1			Session 2			Session 1			Session 2		
		x	y	z	x	y	z	x	y	z	x	y	z
CB	L	**−33**	**−51**	**−33**	−26	−56	−33	**−30**	**−54**	**−35**	2	−65	−17
	R	**26**	**−50**	**−21**	**21**	**−47**	**−21**	**24**	**−50**	**−20**	**20**	**−63**	**−20**

Bold denotes activations corrected for multiple comparisons with FWE $p < 0.05$; non-bold denotes uncorrected activations with a threshold of $p > 0.001$

M1 primary motor cortex, *S1* primary somatosensory cortex, *SMA* supplementary motor area, *CMA* cingulate motor areas, *PMd* dorsal premotor cortex, *PMv* ventral premotor cortex, *SPC* superior parietal cortex, *IPC* inferior parietal cortex, *S2* secondary somatosensory cortex, *CB* cerebellum

movements (Alkadhi et al. 2002; Weiller et al. 1996). By visually inspecting both sessions, the contrast of active movements versus rest showed slightly higher activation than the single contrast using both models. However, activation power increased for the single contrast by including additional movement parameters using model 2, yielding activation patterns largely identical to the contrast with rest.

With respect to reliability, robust activation was elicited consistently with all applied statistical methods and both fMRI models. The size of reliability measures (ICC_{within} and $R_{overlap)}$ on activation maps was in line with the observed activation patterns, with reliability being higher for the contrast with rest and for the single contrast using model 2. To date, only one study tested the reproducibility of brain activation associated with active elbow movements by observing robust reproducible activation in M1 using paired-t-tests (Alkadhi et al. 2002). To our knowledge, the present work is the first study that systematically examines test–retest reliability related to elbow movements. Using a variety of motor tasks, some previous studies reported rather reliable patterns of activation (Alkadhi et al. 2002; Kong et al. 2007; Lee et al. 2010; Yoo et al. 2007). Other studies however, reported large variability across sessions (Kimberley et al. 2008a; Loubinoux et al. 2001; McGonigle et al. 2000). The low reproducibility observed in these investigations probably relies on multiple factors, such as familiarity with the MRI environment and the specific experimental attributes. Diminished attention could also affect brain activation when participants are familiar with the procedure (Loubinoux et al. 2001). Inconsistencies in performance can also induce differences in brain activation, leading to inter-session variability. While some confounding variables, such as familiarity, cannot be completely controlled, differences in task performance can be monitored by MR-compatible devices, which can help to interpret changes in brain activation between sessions. In the present investigation, we used MaRIA in order to keep the experimental settings constant across sessions and monitor the motor performance. Thus, robust activation for active arm movements was assessed successfully. This demonstrates that standardized and well-controlled movement performance improves the reproducibility of brain activation.

The brain network activated by passive elbow movements using MaRIA was comparable to that of active movements and consistent with that reported in a previous study (Weiller et al. 1996). Similar to the findings observed with active movements, the contrast of passive movements with rest showed higher activation than the single contrast using both models. The activation power increased significantly with model 2 through the inclusion of force and dROM in the analysis, leading to largely identical activation patterns to those of the contrasts versus rest. These observations were also in line with the ICC_{within} and $R_{overlap}$ reliability values for activation maps and mainly with $ICC_{between}$ computed on summary statistics, the reliability being higher for contrasts with rest and for single contrasts using model 2. According to the statistical analyses, the reproducibility of brain activation was robust for individual and group activation maps but inconsistent results were found for summary statistics in single ROIs, especially using model 1. Although no study has tested reliability of passive arm movements so far, such tasks had been proposed to elicit brain activation in a more controlled way, as they are independent of the subjects' motor abilities and task requirements (Kocak et al. 2009; Weiller et al. 1996). However, our analyses suggest that, even during passive movements, small differences in task performance do exist in healthy subjects and can potentially affect the reproducibility of activation. Remaining absolutely passive during guided movements is probably quite difficult for healthy subjects. Therefore, we cannot exclude that even with the mechanical device used in our experiment the participants may have squeezed the device's handle differentially or did not follow the movement of the handle in a totally passive way, leading to higher variance across trials in some sessions. This may explain the higher

Table 3 Average $R_{overlap}$ and average ICC_{within} for individual activation maps and $R_{overlap}$ and ICC_{within} for group activation maps of the four contrasts of interest using both models

ROI	Mean Roverlap								Mean ICC							
	Model 1				Model 2				Model 1				Model 2			
	active vs rest	active	passive vs rest	passive	active vs rest	active	passive vs rest	passive	active vs rest	active	passive vs rest	passive	active vs rest	active	passive vs rest	passive
Single subjects																
M1	0.79	0.75	0.74	0.63	0.78	0.77	0.77	0.75	0.90	0.88	0.86	0.80	0.89	0.89	0.87	0.87
S1	0.76	0.69	0.72	0.64	0.72	0.71	0.77	0.74	0.88	0.86	0.86	0.80	0.88	0.87	0.87	0.87
SMA	0.63	0.56	0.49	0.40	0.61	0.61	0.56	0.55	0.80	0.79	0.74	0.63	0.77	0.76	0.75	0.72
PMd	0.76	0.73	0.73	0.64	0.73	0.71	0.76	0.75	0.84	0.83	0.83	0.79	0.83	0.83	0.84	0.83
SPC	0.45	0.44	0.47	0.42	0.46	0.47	0.49	0.49	0.72	0.73	0.64	0.68	0.72	0.71	0.65	0.65
Group																
M1	0.96	0.95	0.84	0.86	0.96	0.96	0.85	0.85	0.97	0.97	0.96	0.95	0.97	0.97	0.96	0.96
S1	0.96	0.92	0.94	0.89	0.94	0.92	0.94	0.92	0.98	0.97	0.93	0.92	0.97	0.97	0.93	0.92
SMA	0.94	0.86	0.94	0.88	0.93	0.91	0.94	0.94	0.93	0.90	0.93	0.94	0.93	0.92	0.92	0.92
PMd	0.95	0.87	0.92	0.83	0.94	0.93	0.92	0.92	0.91	0.92	0.93	0.91	0.93	0.94	0.93	0.92
SPC	0.87	0.80	0.84	0.84	0.86	0.82	0.84	0.79	0.94	0.94	0.90	0.91	0.94	0.94	0.89	0.86

M1 primary motor cortex, *S1* primary somatosensory cortex, *SMA* supplementary motor area, *PMd* dorsal premotor cortex, *SPC* superior parietal cortex. All the ROIs are in the left hemisphere, contralateral to the moving arm

Table 4 $ICC_{between}$ for average and maximum t-values of the four contrasts of interest using both models

ROI	ICC of average t-values								ICC of maximum t-values							
	Model 1				Model 2				Model 1				Model 2			
	active vs rest	active	passive vs rest	passive	active vs rest	active	passive vs rest	passive	active vs rest	active	passive vs rest	passive	active vs rest	active	passive vs rest	passive
M1	0.8*	0.72*	0.65*	0.48*	0.75*	0.72*	0.59*	0.56*	0.82*	0.75*	0.76*	0.56*	0.77*	0.69*	0.73*	0.63*
S1	0.83*	0.77*	0.59*	0.4	0.75*	0.74*	0.58*	0.51*	0.83*	0.81*	0.71*	0.53*	0.82*	0.79*	0.71*	0.62*
SMA	0.67*	0.63*	0.44*	0.44*	0.67*	0.67*	0.47*	0.51*	0.69*	0.72*	0.33	0.36	0.66*	0.71*	0.32	0.37
PMd	0.82*	0.74*	0.4	0.29	0.75*	0.73*	0.38	0.37	0.71*	0.64*	0.63*	0.47*	0.63*	0.6*	0.59*	0.53*
SPC	0.72*	0.65*	0.4	0.26	0.58*	0.53*	0.45*	0.51*	0.81*	0.75*	0.52*	0.38	0.7*	0.66*	0.54*	0.47*

* significant ICC values ($p < 0.05$)

M1 primary motor cortex, *S1* primary somatosensory cortex, *SMA* supplementary motor area, *PMd* dorsal premotor cortex, *SPC* superior parietal cortex. All the ROIs are in the left hemisphere, contralateral to the moving arm

reliability in the active condition, which explicitly required force and joint movements, leading to less variance in performance across trials. Our observations highlight the need for monitoring task performance, both during active and passive movements, and the utility of MRI-compatible robots to address this problem. Furthermore, these findings emphasize the importance of testing the reliability of brain activation patterns, even for passive tasks.

Consistent with previous studies, the ICC_{within} and $R_{overlap}$ values for our group activation maps were highly reproducible in all contrasts and ROIs and were higher than for single subjects (Gountouna et al. 2010; Raemaekers et al. 2007). Across all contrasts of interest and models, $ICC_{between}$ values were lower than for the calculation of ICC_{within}. Lower $ICC_{between}$ values were also reported in several previous studies (Caceres et al. 2009; Raemaekers et al. 2007). A reason for this may have been the low number of subjects usually included in fMRI studies for the $ICC_{between}$ calculation on summary statistics (Caceres et al. 2009). In addition, the low ICC values obtained for passive movements in some ROIs may be attributed to a low level of activation in these areas. For instance, superior parietal cortex was not activated across all subjects using model 1. In contrast, activation in this region was found in all subjects across both sessions using model 2. Overall these new results suggest that activation maps, particularly for group results, are more reliable than summary statistics and that reliability can be improved by enhancing the power of the design, e.g. by increasing the number of trials in the experiment.

As mentioned above for both movement conditions, the higher activation power and reproducibility of brain activation in single contrasts using model 2 may be the consequence of less variance in the performance. Although no differences in mean force and mean dROM were found across repeated measurements, small differences in performance of movements across trials may lead to higher variance in the data and therefore reduced activation power in some subjects. An alternative explanation can be that force and dROM, included as regressors in model 2, may indirectly compensate some motion artifacts potentially correlated to these parameters. Future studies should address this possibility. However, the use of model 2 may be limited when regressors included in the model are strongly correlated with the task (Birn et al. 1999; Johnstone et al. 2006). High correlations may reduce brain activation in some areas. Differences in correlations between sessions may lead to differences in activation and thus, result in misinterpretation of the results. According to earlier publications (Birn et al. 1999; Johnstone et al. 2006), using an event-related design as was done in the present study can overcome this problem. In fact, in our experiment, correlations were very small and constant across both sessions (force: max. mean $r = 0.12$; dROM: max. mean $r = 0.24$). In addition, our results show that the variability can also be reduced by explicitly modeling the rest condition. Such a strategy should also remove variability that cannot be influenced by including motor parameters into the fMRI data analysis, as for example attention changes across sessions. As shown by Specht et al. (2003) attention has an impact on the magnitude of reliability and thus may differently influence passive and active task conditions. The main disadvantage of implementing an additional rest condition is the important increase of scanning time, which is problematic in clinical studies.

In the present investigation, an MR-compatible robot was used to assess arm movement related brain activation while performing active and passive movements. The network activated by the interaction with the robot was consistent with previous studies. The controlled settings reinforced by the device enabled reproducible assessment of brain activation across sessions in single subjects and at group level. Furthermore, quantitative data of the movement performance provided by the device add important information to the analysis. This improved the assessment of brain activation in healthy participants, especially for passive arm movements, by removing variance across trials.

Overall, the results of this study indicate that this device can be used in longitudinal studies to reliably explore brain activation associated with simple arm movements and therefore, is a helpful tool to assess brain reorganization following injury and to monitor rehabilitative interventions in patients with motor impairments. A further application may be the exploration of training induced plasticity in healthy participants to better understand basic mechanisms within the central motor network.

Acknowledgments This work was supported by the National Center of Competence in Research (NCCR) on Neural Plasticity and Repair, launched by the Swiss National Science Foundation (SNF), and by the ETH Research Grant TH-34 06-3 MR-robotics. The authors thank Dr. Birgit Keisker, Prof. Roger Gassert, Prof. Martin Meyer and Dr. Christoph Hollnagel for their helpful advice and comments as well as Dr. Roger Lüchinger for his technical support in the MR center. Special thanks go to all the participants of the study.

References

Alkadhi H, Crelier GR, Boendermaker SH, Golay X, Hepp-Reymond M-C, Kollias SS (2002) Reproducibility of primary motor cortex somatotopy under controlled conditions. Am J Neuroradiol 23(9):1524–1532

Annett M (1970) A classification of hand preference by association analysis. Brit J Psychol 61(3):303–321

Ashburner J (2007) A fast diffeomorphic image registration algorithm. Neuroimage 38(1):95–113

Bennett CM, Miller MB (2010) How reliable are the results from functional magnetic resonance imaging? Ann N Y Acad Sci 1191(1):133–155

Birn RM, Bandettini PA, Cox RW, Shaker R (1999) Event-related fMRI of tasks involving brief motion. Hum Brain Mapp 7(2):106–114

Caceres A, Hall DL, Zelaya FO, Williams SCR, Mehta MA (2009) Measuring fMRI reliability with the intra-class correlation coefficient. Neuroimage 45(3):758–768

Carey LM, Abbott DF, Egan GF, Tochon-Danguy HJ, Donnan GA (2000) The functional neuroanatomy and long-term reproducibility of brain activation associated with a simple finger tapping task in older healthy volunteers: a serial PET study. Neuroimage 11(2):124–144

Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K (2006) The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. Neuroimage 33(2): 430–448

Caspers S, Eickhoff SB, Geyer S, Scheperjans F, Mohlberg H, Zilles K, Amunts K (2008) The human inferior parietal lobule in stereotaxic space. Brain Struct Funct 212(6): 481–495

Cicchetti DV, Sparrow SA (1981) Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 86(2):127–137

Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N (2009) A probabilistic MR atlas of the human cerebellum. Neuroimage 46(1):39–46

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25(4):1325–1335

Eickhoff SB, Schleicher A, Zilles K, Amunts K (2006a) The human parietal operculum. I. Cytoarchitectonic mapping of subdivisions. Cereb Cortex 16(2):254–267

Eickhoff SB, Amunts K, Mohlberg H, Zilles K (2006b) The human parietal operculum. II. Stereotaxic maps and correlation with functional imaging results. Cereb Cortex 16(2):268–279

Eickhoff SB, Paus T, Caspers S, Grosbras M-H, Evans AC, Zilles K, Amunts K (2007) Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36(3): 511–521

Friedman L, Stern H, Brown GG, Mathalon DH, Turner J, Glover GH, Gollub RL, Lauriello J, Lim KO, Cannon T, Greve DN, Bockholt HJ, Belger A, Mueller B, Doty MJ, He J, Wells W, Smyth P, Pieper S, Kim S, Kubicki M, Vangel M, Potkin SG (2008) Test-retest and between-site reliability in a multicenter fMRI study. Hum Brain Mapp 29(8):958–972

Geyer S (2004) The microstructural border between the motor and the cognitive domain in the human cerebral cortex. Adv Anat Embryol Cell Biol 174(I-VIII):1–89

Geyer S, Ledberg A, Schleicher A, Kinomura S, Schormann T, Bürgel U, Klingberg T, Larsson J, Zilles K, Roland PE (1996) Two different areas within the primary motor cortex of man. Nature 382(6594):805–807

Geyer S, Schleicher A, Zilles K (1999) Areas 3a, 3b, and 1 of human primary somatosensory cortex. Neuroimage 10(1):63–83

Geyer S, Schormann T, Mohlberg H, Zilles K (2000) Areas 3a, 3b, and 1 of human primary somatosensory cortex. Neuroimage 11(6):684–696

Gountouna V-E, Job DE, McIntosh AM, Moorhead TWJ, Lymer GKL, Whalley HC, Hall J, Waiter GD, Brennan D, McGonigle DJ, Ahearn TS, Cavanagh J, Condon B, Hadley DM, Marshall I, Murray AD, Steele JD, Wardlaw JM, Lawrie SM (2010) Functional magnetic resonance imaging (fMRI) reproducibility and variance components across visits and scanning sites with a finger tapping task. Neuroimage 49(1):552–560

Grefkes C, Geyer S, Schormann T, Roland P, Zilles K (2001) Human somatosensory area 2: observer-independent cytoarchitectonic mapping, interindividual variability, and population map. Neuroimage 14(3):617–631

Hollnagel C, Brugger M, Vallery H, Wolf P, Dietz V, Kollias S, Riener R (2011) Brain activity during stepping: a novel MRI-compatible device. J Neurosci Meth 201(1):124–130

Johnstone T, Ores Walsh KS, Greischar LL, Alexander AL, Fox AS, Davidson RJ, Oakes TR (2006) Motion correction and the use of motion covariates in multiple-subject fMRI analysis. Hum Brain Mapp 27(10):779–788

Kimberley TJ, Birkholz DD, Hancock RA, VonBank SM, Werth TN (2008a) Reliability of fMRI during a continuous motor task: assessment of analysis techniques. J Neuroimaging 18(1):18–27

Kimberley TJ, Khandekar G, Borich M (2008b) fMRI Reliability in subjects with stroke. Exp Brain Res 186(1):183–190

Kocak M, Ulmer JL, Sahin Ugurel M, Gaggl W, Prost RW (2009) Motor homunculus: passive mapping in healthy volunteers by using functional MR imaging–initial results. Radiology 251(2):485–492

Kong J, Gollub RL, Webb JM, Kong J-T, Vangel MG, Kwong K (2007) Test-retest study of fMRI signal change evoked by electroacupuncture stimulation. Neuroimage 34(3):1171–1181

Lee JN, Hsu EW, Rashkin E, Thatcher JW, Kreitschitz S, Gale P, Healy L, Marchand WR (2010) Reliability of fMRI motor tasks in structures of the corticostriatal circuitry: implications for future studies and circuit function. Neuromage 49(2):1282–1288

Loubinoux I, Carel C, Alary F, Boulanouar K, Viallard G, Manelfe C, Rascol O, Celsis P, Chollet F (2001) Within-session and between-session reproducibility of cerebral sensorimotor activation: a test–retest effect evidenced with functional magnetic resonance imaging. J Cereb Blood Flow Metab 21(5):592–607

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19(3):1233–1239

Mayka MA, Corcos DM, Leurgans SE, Vaillancourt DE (2006) Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. Neuroimage 31(4):1453–1474

McGonigle DJ, Howseman AM, Athwal BS, Friston KJ, Frackowiak RS, Holmes AP (2000) Variability in fMRI: an examination of intersession differences. Neuroimage 11(6 Pt 1):708–734

McGregor KM, Carpenter H, Kleim E, Sudhyadhom A, White KD, Butler AJ, Kleim J, Crosson B (2012) Motor map reliability and aging: a TMS/fMRI study. Exp Brain Res 219(1):97–106

Raemaekers M, Vink M, Zandbelt B, van Wezel RJA, Kahn RS, Ramsey NF (2007) Test–retest reliability of fMRI activation during prosaccades and antisaccades. Neuroimage 36(3):532–542

Rombouts SA, Barkhof F, Hoogenraad FG, Sprenger M, Scheltens P (1998) Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging. Magn Reson Imaging 16(2):105–113

Scheperjans F, Eickhoff SB, Hömke L, Mohlberg H, Hermann K, Amunts K, Zilles K (2008a) Probabilistic maps, morphometry, and variability of cytoarchitectonic areas in the human superior parietal cortex. Cereb Cortex 18(9):2141–2157

Scheperjans F, Hermann K, Eickhoff SB, Amunts K, Schleicher A, Zilles K (2008b) Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. Cereb Cortex 18(4):846–867

Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86(2):420–428

Specht K, Willmes K, Shah NJ, Jäncke L (2003) Assessment of reliability in functional imaging studies. J Magn Reson Imaging 17(4):463–471

Tsekos NV, Khanicheh A, Christoforou E, Mavroidis C (2007) Magnetic resonance-compatible robotic and mechatronics systems for image-guided interventions and rehabilitation: a review study. Ann Rev Biomed Eng 9:351–387

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Étard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289

Weiller C, Jüptner M, Fellows S, Rijntjes M, Leonhardt G, Kiebel S, Müller S, Diener HC, Thilmann AF (1996) Brain representation of active and passive movements. Neuroimage 4(2):105–110

Yoo S–S, O'Leary HM, Lee J-H, Chen N-K, Panych LP, Jolesz FA (2007) Reproducibility of trial-based functional MRI on motor imagery. Int J Neurosci 117(2):215–227

Yu N, Hollnagel C, Blickenstorfer A, Kollias SS, Riener R (2008) Comparison of MRI-compatible mechatronic systems with hydrodynamic and pneumatic actuation. IEEE Asme T Mech 13(3):268–277

Yu N, Hollnagel C, Wolf P, Murr W, Blickenstorfer A, Kollias S, Riener R (2009) Tracking and analysis of human head motion during guided fMRI motor tasks. IEEE ICORR 2009:588–593

Yu N, Estévez N, Hepp-Reymond M-C, Kollias SS, Riener R (2011) fMRI Assessment of upper extremity related brain activation with an MRI-compatible manipulandum. Int J Comput Assist Radiol Surg 6(3):447–455