Appendix e-1
Supplemental Information

Study medication packing, blinding and random allocation sequences:
The study drug and placebo (Salinex® saline nasal spray) were repackaged by the research pharmacy at our institution (London Health Sciences Centre) into matching sterile glass bottles fitted with spray nozzles using standard aseptic techniques and were refrigerated until use (within 30 days). All random allocation sequences were devised and generated by the research pharmacy. For the first three dose cohorts, a random number between 1 and 5 was generated and assigned to receive the placebo. The gender of the fifth patient enrolled was left unassigned until complete enrollment of patients 1-4 for each dose. The research pharmacy then communicated the gender required for patient #5 to the study coordinator to ensure that n=2 males and n=2 females received oxytocin at each dose. For the final cohort of 8 patients randomized in a 1:1 ratio to 72 IU or placebo, random allocation was done in two blocks of four (2 active and 2 placebo per block). Sequential numbering of medication containers was used throughout the study to conceal the allocation sequence.

Estimated duration of action of intranasal oxytocin:
Much of the initial work evaluating oxytocin administration in humans was based on extrapolation from pharmacokinetics of intranasal vasopressin administration, as AVP differs from oxytocin by only 2 amino acid, which demonstrated increased levels of AVP in the CSF beginning 10 minutes after intranasal administration, and were elevated still at 80 and 120 minutes. [1] CSF levels of oxytocin obtained by lumbar puncture have been measured in a small number of health adults [2] and showed significant increases (64%) at 75 minutes. These findings are consistent with those in rhesus macques where a demonstrated ~2.5 fold increase in CSF oxytocin levels was ~35 min post intranasal administration of 25 IU oxytocin. [3] The duration of CSF elevations of oxytocin is not yet characterized in humans as levels have typically remained elevated after the last CSF measurement. As CSF volumes are renewed in ~5 hours, [2] we elected to use morning and afternoon dosing to maximize the awake hours in which oxytocin elevations may be present in patients. The present study design we did not include CSF measurements due to concerns that this would prohibit sufficient patient recruitment.

Subject Inclusion Criteria:

-Age 30-80
-Meets Neary criteria* for diagnosis of probable frontotemporal dementia (bvFTD or semantic dementia with significant behavioural features including empathy deficits)
-Neuroimaging (CT, MRI or SPECT scan) supports diagnosis of Frontotemporal dementia
-Provides written informed consent and has a caregiver or legally acceptable representative who provides written informed consent.

Subject Exclusion Criteria.

-Has a history of a myocardial infarction within the last two years or congestive heart failure.
-Current uncontrolled hypertension
-Current bradycardia (rate <50 bpm) or tachycardia (rate > 100 bpm)

-Current hyponatremia
-Current use of prostaglandin medications
-Females who are pregnant or breastfeeding
-Use of any investigational or experimental drug or device within the last 60 days prior to screening or within 5 half-lives of the experimental drug, whichever is longer.

Main Efficacy Measure

The Neuropsychiatric Inventory. [4] The Neuropsychiatric Inventory is a caregiver response measure completed via interview reflecting behavioural and personality changes in the patient. Items include delusions, hallucinations, agitation/aggression, depression, anxiety, elation, apathy, disinhibition, irritability, as well as aberrant motor behaviour, sleep and appetite. Each item is rated on dimensions of behaviour frequency, severity and distress. Total scores were calculated by multiplying frequency by severity scores for each item and then summing the domain scores; higher scores indicate greater behavioural impairment. [5] The NPI was designated as the primary efficacy measure based on prior findings of a significant improvement in the NPI total scores in patients with FTD following a single dose of oxytocin compared to placebo.[6]

Secondary Efficacy Measures

Interpersonal Reactivity Index.[7] The interpersonal reactivity index (IRI is a 28-item questionnaire answered on a 5-point Likert scale which ranges from "Does not describe me well" to "Describes me very well". The measure has 4 subscales: perspective taking, fantasy, empathic concern and personal distress. Perspective taking is defined as the tendency to spontaneously adopt the psychological point of view of others. Fantasy evaluates the tendencies to imagine the feelings and actions of characters in books, movies and plays. Personal distress measures feelings of personal anxiety and unease that is self-oriented in tense interpersonal settings. Lastly, empathic concern indexes feelings of sympathy and concern for others in unfortunate circumstances. An increase in these scores indicates that subscale is more descriptive of that individual. The IRI has been used previously in patients with FTD, who demonstrate deficits in the Empathic concern and Perspective taking subscales; [8-12] which are thought to relate to emotional and cognitive empathy respectively. Based on prior reports showing oxytocin augmented emotional empathy but not cognitive empathy,[13] we predicted the empathic concern score would be improved following oxytocin administration.

Clinical Global Impression.[14] The Clinical Global Impression scale is a 3-item observer-rated scale that measures illness severity (CGIS), global improvement or change (CGIC) and therapeutic response. The CGI is rated on a 7-point scale, with the severity of illness scale allowing a range of responses from 1 (normal) through to 7 (amongst the most severely ill patients). CGI-C scores range from 1 (very much improved) through to 7 (very much worse). Treatment response ratings take account of both therapeutic efficacy and treatment-related adverse events and range from 0 (marked improvement and no side-effects) to 4 (unchanged or worse and side-effects outweigh the therapeutic effects).

The Frontal Behavioural Inventory.[15] The Frontal Behavioural Inventory is a caregiver response inventory completed via interview, reflecting behavioural change in patients with FTD. Nine

sub-items from this scale related to social cognition or repetitive behaviours hypothesized to be affected by oxytocin were administered: apathy, indifference/emotional flatness, perseverations and obsessions, inappropriateness, excessive jocularity, impulsivity/poor judgment, irritability, aggression and hypersexuality. Higher scores are indicative of greater behavioural impairment; the highest possible total score on these sub-items was 27.

The Multifaceted Empathy Task (MET).[16] The MET is a computerized performance-based empathy measure that assesses cognitive and emotional empathy in a dissociable way. The MET has been used to demonstrate changes in empathy following oxytocin administered to healthy adults, and has recently been used to characterize empathy deficits in FTD.[17] The MET also includes both negative and positive stimuli, allowing for the valence-based discrimination of empathic responses. Participants are presented with 23 pairs of realistic emotionally-charged pictures: a context-only image that is presented first, and an image of the same context with a person or people embedded within it that appears later. For the context-only images, participants provide valence and arousal ratings using 9-point Likert scales. Cognitive empathy is measured by asking participants to infer the mental states of individuals in the photos from four options. Emotional empathy is measured by obtaining participants' explicit and implicit ratings of their emotional experiences in response to the pictures on 9-point Likert scales. As above, in healthy adults oxytocin administration augmented emotional empathy but not cognitive empathy performance on the MET,[13] we predicted specifically that the emotional empathy scores would be improved following oxytocin administration.

Frontotemporal Dementia Rating Scale.[18] The FRS is a severity rating tool for patients with frontotemporal dementia. The FRS is based on a caregiver interview and assesses patient's functional abilities including reports on daily performance related to behaviours, outings and shopping, household chores, finances, medications, meal preparation, mobility and self-care.

Frontotemporal-Dementia modified Clinical Dementia Rating Scale.[19] The CDR structured interview collects information from both the participant and the caregiver that was developed primarily for the assessment of Alzheimer's disease but has been modified to assess functioning in individuals with FTD. Compared to the original version of the CDR, the modified version takes into account behavior and personality changes as well as language impairments that often accompany FTD. Other domains measured are: memory, orientation, judgment and orientation, community affairs, home and hobbies, and personal care. Scores range from 0 (no cognitive impairment) to 3 (severe cognitive impairment).

Baseline Symptoms and Adverse Events Checklist. The Baseline Symptoms and Adverse Events Checklist consists of 29 items that ask whether that symptom is "Absent" or "Present". For each item listed, if symptoms are present the severity, chronicity, and description of the symptom and onset/cessation date are recorded. Item 29 inquires about any other symptoms that may not have been covered in the questionnaire to ensure all adverse events are recorded.

Apathy Evaluation Scale.[20] The apathy evaluation scale consists of 18 caregiver rated items to evaluate apathy as reflected by behavior, cognition and emotion as scored by the caregiver. The

items are rated on a 4-point Likert scale from 1 (not at all characteristic) to 4 (very characteristic). Higher scores indicate higher levels of apathy.

Statistical Analysis:
Statistical analyses were performed in the Statistical Package for the Social Sciences (SPSS, Version 21).

Supplemental Results:

Multifaceted Empathy Test: Twenty-one patients were able to complete the MET at both visits (n=7 patients receiving placebo and n=14 patients receiving oxytocin). Of the 2 patients unable to complete the task, one was unable to understand the rating system and the second patient refused to complete the task on the Day 7 visit.

Apathy Evaluation Scale: As the planned interim analysis conducted after the first three cohort dose groups were completed had indicated improvement on the NPI apathy subdomain in the oxytocin treatment group, the AES was administered at baseline and follow up for the final 8 participants (n=4 randomized to 72 IU and n=4 to placebo). Given the sample size, no statistical tests were conducted.

Power Table for Sample Size Calculations

Sample Size total	**P value**	**Power†**
30	0.05	0.67
36	0.05	0.75
40	0.05	0.79
44	**0.05**	**0.83**
50	0.05	0.88
54	0.05	0.9
60	0.05	0.93
66	0.05	0.93
Power calculations performed using the statistical power calculator for parallel group clinical trials available at http://hedwig.mgh.harvard.edu/sample_size/js/js_parallel_quant.html		
†Power calculations based on two group design (placebo vs. oxytocin), based on standard deviation in NPI apathy domain scores of 3.3, and minimal detectable difference that is estimated to be clinically significant of 3 point change.		
Sample Size total	**P value**	**Power***
30	0.05	0.91
36	0.05	0.95
40	0.05	0.97
44	0.05	0.98
50	0.05	0.99
54	0.05	0.99
60	0.05	0.99
66	0.05	0.99
*Power calculations based on two group design (placebo vs. oxytocin), based on standard deviation in FBI apathy scores of 0.8, and minimal detectable difference that is estimated to be clinically significant of 1 point change.		

References

1. Born J, Lange T, Kern W, McGregor GP, Bickel U, Fehm HL. Sniffing neuropeptides: a transnasal approach to the human brain. Nat Neurosci 2002;5:514-516.

2. Striepens N, Kendrick KM, Hanking V, et al. Elevated cerebrospinal fluid and blood concentrations of oxytocin following its intranasal administration in humans. Sci Rep 2013;3:3440.

3. Chang SW, Barter JW, Ebitz RB, Watson KK, Platt ML. Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (Macaca mulatta). Proc Natl Acad Sci U S A 2012;109:959-964.

4. Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997;48:S10-16.

5. Connor DJ, Sabbagh MN, Cummings JL. Comment on administration and scoring of the Neuropsychiatric Inventory in clinical trials. Alzheimers Dement 2008;4:390-394.

6. Jesso S, Morlog D, Ross S, et al. The effects of oxytocin on social cognition and behaviour in frontotemporal dementia. Brain 2011;134:2493-2501.

7. Davis MH. A multidimensional approach to individual differences in empathy. In: JSAS Catalog of Selected Documents in Psychology, 1980.

8. Lough S, Kipps CM, Treise C, Watson P, Blair JR, Hodges JR. Social reasoning, emotion and empathy in frontotemporal dementia. Neuropsychologia 2006;44:950-958.

9. Rankin KP, Gorno-Tempini ML, Allison SC, et al. Structural anatomy of empathy in neurodegenerative disease. Brain 2006;129:2945-2956.

10. Eslinger PJ, Moore P, Anderson C, Grossman M. Social cognition, executive functioning, and neuroimaging correlates of empathic deficits in frontotemporal dementia. Journal of Neuropsychiatry and Clinical Neurosciences 2011;23:74-82.

11. Hsieh S, Irish M, Daveson N, Hodges JR, Piguet O. When One Loses Empathy: Its Effect on Carers of Patients With Dementia. Journal of Geriatric Psychiatry and Neurology 2013;26:174-184.

12. Fernandez-Duque D, Hodges SD, Baird JA, Black SE. Empathy in frontotemporal dementia and Alzheimer's disease. J Clin Exp Neuropsychol 2010;32:289-298.

13. Hurlemann R, Patin A, Onur OA, et al. Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. J Neurosci 2010;30:4999-5007.

14. Schneider LS, Olin JT, Doody RS, et al. Validity and reliability of the Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11 Suppl 2:S22-32.

15. Kertesz A, Davidson W, Fox H. Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. Can J Neurol Sci 1997;24:29-36.

16. Dziobek I, Rogers K, Fleck S, et al. Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). J Autism Dev Disord 2008;38:464-473.

17. Oliver L, MacKinley J, Coleman K, et al. Parsing cognitive and emotional empathy deficits for negative and positive stimuli in frontotemporal dementia. Neuropsycholgia under revision.

18. Mioshi E, Hsieh S, Savage S, Hornberger M, Hodges JR. Clinical staging and disease progression in frontotemporal dementia. Neurology 2010;74:1591-1597.

19. Knopman DS, Kramer JH, Boeve BF, et al. Development of methodology for conducting clinical trials in frontotemporal lobar degeneration. Brain 2008;131:2957-2968.

20. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. Psychiatry Res 1991;38:143-162.

Figure e-1. Study Elements

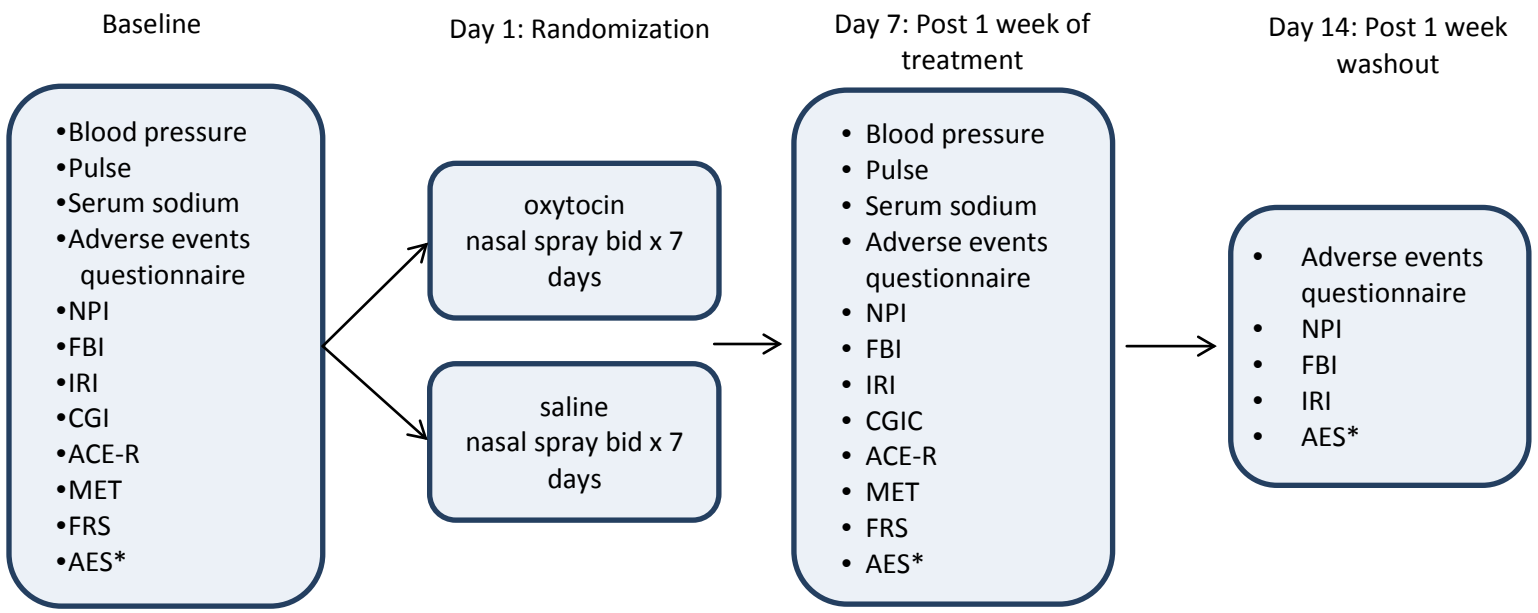

Figure e-1. Study Elements. Safety assessments including blood pressure, pulse, serum sodium levels, adverse events questionnaire, caregiver ratings of patient's behaviours (NPI, FBI, IRI) and disease severity (CDR-FTLD, FRS), physician-completed severity index (CGI), and patient's performance on an empathy measures (Multi-faceted Empathy Test) were completed at baseline and after 7 days of treatment. Day 14 assessments (after 1 week of medication washout) were completed via telephone interview with caregivers to assess effects of medication discontinuation. ACE-R= Addenbrook Cognitive Exam-Revised, MET= Multi-faceted Empathy Test,

NPI= Neuropsychiatric Inventory, CGI= Clinician's Global Impression, CGIC= Clinician's Global Impression of Change, FBI= Frontal Behavioural Inventory, IRI=Interpersonal Reactivity Index, FRS= Frontotemporal Dementia Rating Scale, AES=Apathy Evaluation Scale. * Note, AES was added to study for last n=8 patients enrolled to further examine measures indexing apathy.

Figure e-2. Dose Escalation Model

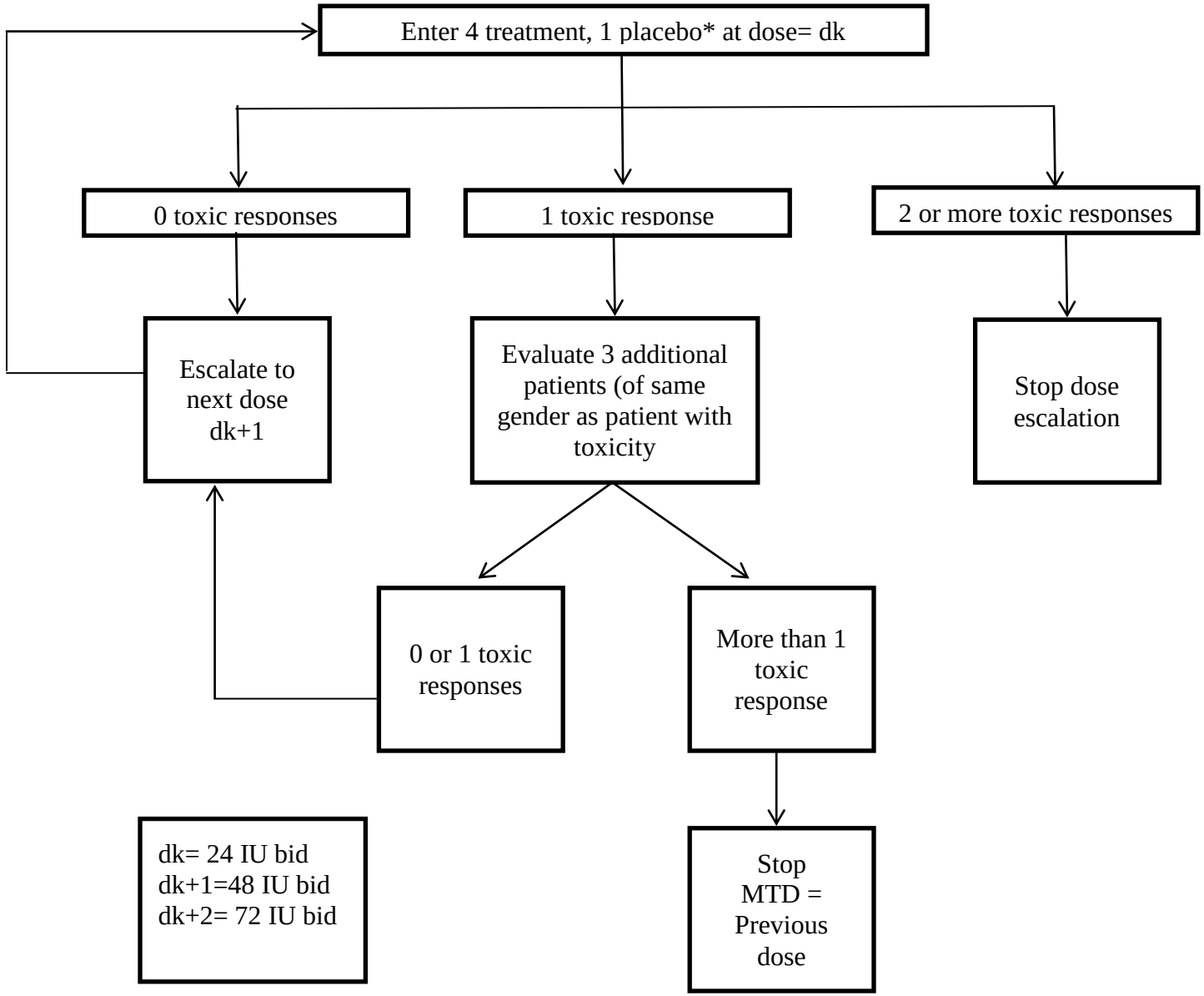

Figure e-2. Dose escalation model. This trial design was modeled on the classical "3+3" design for dose finding studies,[15,16] with two modifications. As differential effects of oxytocin may be expected for males and females, rather than a total of three patients per arm, to ensure that at least 2 males and 2 females received oxytocin at each dose prior to the dose escalation we enrolled a total of 5 patients into each dosage arm. Two males and two females were randomized to study drug or placebo. Following completion of all the study measures by these four participants, the research pharmacy designated the gender of the fifth participant to be enrolled in each dose group. When the maximum tolerated dose (MTD) was identified, 8 additional patients were enrolled (4 randomized to placebo and 4 to oxytocin treatment) so that a minimum of 8 patients received the MTD.

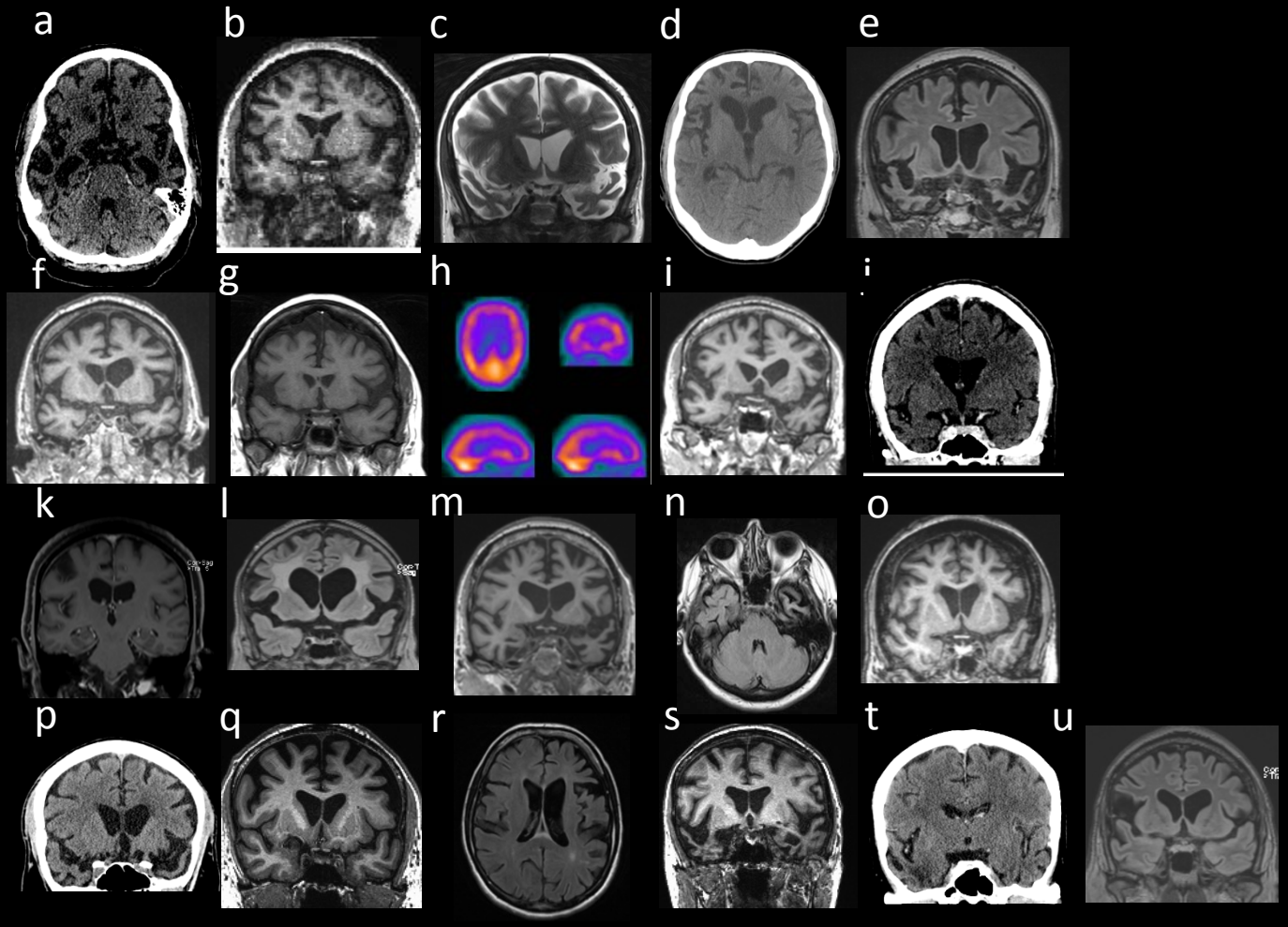

Figure e-3. Clinical diagnostic imaging of study participants supporting diagnosis of probable FTD (see table e-1 for details).

Table e-1. Clinical imaging findings for study participants.

Image	Dose Group	Diagnosis	Neuroimaging study	Neuroimaging findings
a	24	bvFTD	MRI	bitemporal and frontal atrophy
b	24	bvFTD	MRI	L>R temporal and frontal atrophy
c	24	SD+bvFTD	MRI	bitemporal atrophy, L>R
d	24	bvFTD	CT	bifrontal atrophy, R>L
e	48	bvFTD	MRI	bitemporal atrophy, L>R (left handed)
f	48	bvFTD	MRI	bifrontal and temporal atrophy
g	48	bvFTD	MRI	bifrontal atrophy, parietal atrophy
h	48	bvFTD	SPECT	R>L temporal and inferior frontal hypoperfusion
i	72	bvFTD$^{\pm}$	MRI	bifrontal, temporal and parietal atrophy
j	72	bvFTD	CT	bitemporal and frontal atrophy R>L
k	72	bvFTD	MRI	right frontal atrophy
l	72	bvFTD	MRI	bifrontal atrophy
m	72	bvFTD$^{\pm}$	MRI	bifrontal and temporal atrophy
n	72	SD+bvFTD	MRI	bitemporal atrophy, L>R
n/a*	72	bvFTD	MRI	bifrontal atrophy
n/a*	72	bvFTD	SPECT	right temporal hypoperfusion
o	placebo	bvFTD	MRI	bifrontal atrophy
p	placebo	bvFTD	MRI	bifrontal and temporal atrophy
q	placebo	bvFTD+PNFA	MRI	frontal atrophy L>R
r	placebo	bvFTD+PNFA	MRI	frontal atropy L>R
s	placebo	SD+bvFTD	MRI	bitemporal atrophy R>L
t	placebo	bvFTD	CT	bitemporal atrophy R>L
u	placebo	bvFTD	MRI	bifrontal atrophy R>L

R= right, L= left. SD= semantic dementia, bvFTD= behavoural variant FTD, PNFA= progressive nonfluent aphasia. *Images not available for publication from one patient referred from another centre, and 1 patient with SPECT scan from main centre. $^{\pm}$ diagnosis also supported by expanded repeats at *C9ORF72*.

Table e-2. Mean differences in safety measures from baseline to Day 7

		Oxytocin by Dose Subgroup							
	Placebo	**Combined**	**p-value**	**24 IU**	**p-value**	**48 IU**	**p-value**	**72 IU**	**p-value**
Heart Rate (beats per minute)	-2·86 (-6·8 to 1·1)	-1·75 (-5·8 to 2·4)	0·52	-1·50 (-15·1 to 12·1)	0·34	-4·83 (-7·9 to 0·1)	0·70	-0·75 (-8·7 to 7·2)	0·45
Systolic Blood Pressure (mmHg)	5·29 (-2·9 to 13·5)	-3·25 (-7·2 to 0·7)	0·04	-1·00 (-8·6 to 6·6)	0·25	0·25 (-9·3 to 9·8)	0·22	-6·13 (13·1 to 0·9)	0·04
Diastolic Blood Pressure (mmHg)	2·43 (-5·4 to 10·3)	-3·37 (-6·2 to -0·6)	0·11	-3·50 (-12·2 to 5·3)	0·26	-5·25 (-6·7 to -3·7)	0·18	-2·38 (-7·8 to 3·0)	0·22
Sodium Levels (mEq/L)	-0·43 (-2·0 to 1·2)	-0·43 (-1·2 to -0·3)	0·97	-1·50 (-4·3 to 1·3)	0·34	0 (-1·8 to 1·8)	0·70	-0·13 (1·2 to 0·9)	0·77

Data are mean (95% confidence inter