

融合画像としての FDG および FAMT-PET/CT による癌の局在・進展度診断と半導体コンプトンカメラへの発展性について

研究分担者 対馬 義人 群馬大学大学院医学系研究科 教授

研究要旨：コンプトンカメラは、 γ 線のコンプトン散乱を利用し、 γ 線の発生位置とエネルギースペクトルを測量することができる装置であり、宇宙実験物理学の分野で利用される。これを医療に応用する場合、革新的な小型簡易型の PET/SPECT 装置の完成が期待される。その応用として、訪問診療、災害時医療など広範な領域での臨床応用が考えられる。一方、本学の実際の核医学診療でも、コンプトンカメラの利用が期待される。

難治性の癌の重粒子線治療では、 ^{12}C 粒子線をターゲットに照射すると入射粒子とターゲット中の粒子との相互作用によって微量に生じる ^{11}C の分布を PET 画像で評価可能であり、確実な照射野の確認、治療域の決定が可能となる。この画像をコンプトンカメラで評価する事も可能かもしれない。また、現在進行中の核医学診療でも PET/CT や PET/MRI などの融合画像が正確な腫瘍の局在診断とそれに基づく治療計画策定に重要である。

本学では、FDG および FAMT-PET/CT を用いて、癌の局在診断と進展診断を日常臨床にて行っている。本報告では、現時点までのそれらの成果を報告する。

A. 研究目的

高齢化社会の到来とともに、癌の罹患率と死亡率も増加の一途をたどっている。核医学診療は、癌の診断およびその原理を利用した治療が可能であり、ラジオアイソトープ (RI) が使われる。

本学では、陽電子放出核種 (ポジトロン核種) を用いて生体の代謝や臓器血流を定量評価する PET (positron emission tomography) 検査は 20 年以上臨床応用されている。現有する PET と CT の複合機 (SIEMENS・GE 社製) は、PET 画像と CT 画像の融合画像が作成でき、病変部のより正確な診断と位置の特定が可能になった。そし

て、ブドウ糖やアミノ酸、核酸の標識化合物を代謝マーカーとして用いた悪性腫瘍などの分子イメージングとして臨床・研究に活用されている。

近年は、PET/CT に加え、PET と MRI 画像の融合による PET/MRI 検査も可能となり、融合画像のもつ、がん診断、治療へのインパクトは非常に大きいと言える。

世界的に最も汎用されている ^{18}F -FDG は、癌細胞表面に過剰発現されているブドウ糖担体や、細胞内のリン酸化酵素であるヘキソキナーゼの活性と関連し癌細胞に取り込まれる。多くの腫瘍細胞は糖代謝が活発なため ^{18}F -FDG は強く集積し、病変と正常組織

との高いコントラストにより高精度の診断が可能である。しかし頭頸部などにおいては、領域が狭小であり、筋肉や炎症、良性腫瘍など、種々の非特異的集積により、 ^{18}F -FDGを用いた時に正確な評価が困難な場合がある。

一方、L-3-[fluorine-18] fluoro-alpha-methyl tyrosine (以下 ^{18}F -FAMT と略) は、癌細胞に選択的に過剰発現している L 型中性アミノ酸の細胞内への輸送に係わる、細胞表面の担体である LAT1 (L-type amino acid transporter 1) を介して腫瘍細胞内に取り込まれる。これは、我々の施設でアミノ酸代謝の評価を目的に PET 検査に用いるものとして開発され臨床応用されている核種である。

本学では、顎口腔腫瘍に対して多くの FAMT-PET を行っている。顎口腔腫瘍の術前画像診断により悪性腫瘍の進展範囲を正確に把握し、健常組織の過剰な切除を防ぐことは、術後の整容や咀嚼・嚥下・発語機能改善に寄与し、社会的生命予後の向上につながる。

一方、骨転移は担癌患者にて頻度の高い疾患であり、その有無によってステージングが決まる。また、骨転移によって起こり得る症状が QOL 低下をきたすため、早期に骨転移を診断し、治療効果を判定することが重要である。骨転移の評価としては、単純 X 線、骨シンチグラフィ、FDG-PET で評価が行われる。最近では DWIBS という MRI の拡散強調画像による全身撮像法や PET 検査の中でも FDG とは異なる NaF や ^{11}C メチオニンがあるが、まだごく限られた施設でしか行えない。そこで、我々の施設ではメチオニンと同じアミノ酸製剤である ^{18}F -FAMT を利用

した PET 検査を行っている。

悪性腫瘍での骨転移病変の広がりや、骨シンチ、FDG、FAMT でそれぞれ集積機序が異なることから、異なる像を呈する。複数の核種で検査を同時に行うことができればそれぞれの核種の分布から得られる情報を総合して、より正確な腫瘍細胞の進展範囲の診断が可能であり、さらには、放射線治療におけるより正確な照射野の設定や、外科的切除におけるより適切な切除範囲の決定が可能となる。

この様な診断が、コンプトンカメラを用いることによってより正確に可能となる可能性があり、今年度は複数の核種での病変進展範囲の違いをターゲットに評価を行ってきている。

B. 研究方法

骨転移の多核種による分布の違いに関する評価にターゲットを絞って、その診断において、FDG-PET を基準とした FAMT の診断能力はどうか？そしてその集積に特徴はあるのか？今後の診断有用性はどうか？これらに対して検討を行った。

2010 年 8 月から 2011 年 10 月にかけて、FDG と FAMT-PET をそれぞれ一か月以内に撮影した、骨転移を合併する悪性疾患の 21 症例 72 部位を対象に後ろ向きに検討を行った。年齢は 55 歳から 81 歳で、男性 15 例、女性 6 例である。

FDG、FAMT-PET とともに 6 時間の絶食後、5MBq/kg を静注し、60 分後に PET/CT の撮像を行った。集積自体の評価については、SUV 値という半定量値の最大値である SUV max で行った。定性的評価として、他のモダリティや臨床的に診断された骨転移に対して、

FDG-PET の集積部位を同定し、FAMT-PET で同部位の集積状態を評価した。そして、CT でその部位が造骨性なのか溶骨性なのか形態的評価を行った。また、FDG と FAMT の相関関係を調べ、集積に有意差があるのか、t 検定を行った。

一方、顎口腔腫瘍の検討では、歯肉扁平上皮癌の顎骨浸潤および腫瘍進展範囲における各種画像診断機器の有用性について比較検討した。対象は 2008 年 4 月から 2013 年 9 月まで、群馬大学口腔外科を受診した 40 症例である。

まず、MRI、PET による顎骨浸潤の診断能について視覚的に評価した。顎骨浸潤のスコア分類については、Yasser らの分類に基づいて、スコア 2 以下を顎骨浸潤がないもの、3 以上を顎骨浸潤があるものとして分類した。次に、病理組織標本と PET 集積の体積を計測し、腫瘍の進展範囲として半定量的に評価した。PET 集積の測定方法として、PETVCAR という GE 社のソフトウェアを使用した。PET-VCAR を用いた PET 画像上で計測した腫瘍の推定体積と、切除標本から概算した腫瘍体積とを比較した。その際のしきい値は、FDG で 3、FAMT で 1.4 以上とし、切除標本の腫瘍体積はスライドの式より概算した。

C. 研究結果

全骨転移において、全体的に FAMT は FDG と比べて集積が弱い、FAMT によって骨転移を指摘できたのは 72 部位中 63 部位の 87.5%であった。また、変形性変化を示す 38 部位中、FAMT で集積があったのは一例のみで、特異度は 97.4%であった。FAMT と FDG の SUV max の相関関係では、全症例対象で、

$r=0.26$ で弱いものの有意な相関があった。原発腫瘍別で FAMT と FDG の相関関係を見ると、肺癌が最も相関関係が高い結果となった。

原発腫瘍別の FAMT と FDG の集積比を調べると、食道癌での比が最も高く、FDG により近い集積を示す結果となった。組織型別では、腺癌、扁平上皮癌共に FDG と FAMT の SUV max との相関関係に有意なものが得られた。そして、FAMT と FDG の比から両者の特徴を調べたところ、それぞれに有意差を認め、腺癌に比べて扁平上皮癌の方が、比は高いという結果となった。

造骨性と溶骨性での FAMT と FDG の集積は、FAMT では造骨性、溶骨性に有意差はなく、FDG では有意差を認めた。FDG は溶骨性腫瘍に強く集積すると一般的に言われているが、今回の結果はそれを裏付けるものであり、また FDG では造骨性、溶骨性といった骨転移様式で集積に差がある一方、FAMT では差がないことがわかった。

顎口腔腫瘍の症例を対象にした検討では、骨髄浸潤の有無について、感度が最も高かったのは、FDG であった。しかし、下顎歯肉癌は視診にて病変部位は確認できることから、画像診断で最も重要視される真陰性は FAMT 13 例に対し FDG はわずか 2 症例であった。この結果より、特異度および正診率では FAMT がより高く、他の画像検査を補完する可能性が示唆された。病理組織標本と PET の集積予測体積についてそれぞれ回帰分析を行った結果、FDG の相関係数は $r=0.47$ と弱い相関に対し、FAMT は $r=0.72$ と強い相関を認めた。また、寄与率においても FDG と比較し FAMT でより高い値を示した。これら病理組織標本との強い相関からも、FAMT は

適正な顎骨切除範囲や治療方針を予測するための指標となる可能性を認めた。

D. 考察

FAMT-PET でも多くの骨転移病変は検出可能であるが、原発腫瘍や組織型、骨転移の性状により相関関係や比に相違を認めるため、現時点ではスタンダードとなっている FDG と相補的に骨転移診断に利用するのが望ましいと考えられる。

FAMT 集積が相対的に強い食道腫瘍や扁平上皮癌の腫瘍では、骨転移診断における有用性が期待される。FAMT の集積の特徴として悪性腫瘍に特異的に集積すること、生理的集積が少ないことが挙げられる。したがって背景と病変とのコントラストが良好であり、病変の正確な判断が可能となる。

FAMT は FDG と比べて腫瘍の進展範囲を正確に反映している可能性があり、FDG よりも更に正確な進展範囲の診断や治療効果判定が行えるのではないかとということが示唆される。

放射線治療や重粒子線治療など治療後評価に造影 CT や FDG-PET を行っているが、二次性変化か残存腫瘍か判断が難しい場合が多いため、悪性腫瘍に特異的に集積する特徴と併せて、そのような症例に FAMT は大変有用である。

顎骨への腫瘍浸潤評価については、これまで歯科用コンビーム CT や MRI など報告されているが、骨髄浸潤の予測は CT 単独では困難である。また、骨髄内への進展範囲は、MR やシンチなどでは炎症の波及も構築画像に影響するため、腫瘍の過大評価が懸念される。

本研究により、悪性腫瘍に特異的に集積

する FAMT-PET の有用性が示唆された。

E. 結論

FAMT による骨転移の診断能力は優れていると言え、FAMT は扁平上皮がんの腫瘍に対して、更に診断価値を高める放射性医薬品である。FAMT は骨転移タイプに左右されない診断能力があり、他のモダリティに対して優位な点である。

一方、顎骨浸潤の進展範囲の診断においても FDG-PET や MRI とくらべて、FAMT-PET でのより正確な顎骨浸潤評価が可能であることが確認できた。

今後、これらの結果を踏まえ、複数の核種の同時投与により、コンプトンカメラでの検査が可能になることが期待される。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

Morita M, Higuchi T, Achmad A, Tokue A, Arisaka Y, Tsushima Y. Complementary roles of tumour specific PET tracer ^{18}F -FAMT to ^{18}F -FDG PET/CT for the assessment of bone metastasis. *Eur J Nucl Med Mol Imaging.* 2013; 40(11):1672-1681.

Kim M, Higuchi T, Arisaka Y, Achmad A, Tokue A, Tominaga H, Miyashita G, Miyazaki H, Negishi A, Yokoo S, Tsushima Y. Clinical significance of ^{18}F - α -methyl tyrosine PET/CT for the detection of bone marrow invasion in

patients with oral squamous cell carcinoma: comparison with ^{18}F -FDG PET/CT and MRI. Ann Nucl Med. 2013; 27(5):423-430.

2. 学会発表

金舞, 樋口徹也, 有坂有紀子, 徳江梓, 菅原茂耕, 宮崎英隆, 根岸明秀, 横尾聡, 対馬義人: 下顎歯肉扁平上皮癌の顎骨浸潤検出における FMT-PET/CT の臨床的意義: FDG - PET/CT と MRI との比較. 第 53 回日本核医学会学術総会 11. 8-10, 福岡

金舞, 樋口徹也, 有坂有紀子, 宮崎英隆, 根岸明秀, 横尾聡, 対馬義人: 口腔腫瘍における FMT-PET/CT の早期画像の有用性について. PET サマーセミナー2013 in 加賀百万石 8. 23-25, 石川

金舞, 樋口徹也, 有坂有紀子, 徳江梓, 森田元穂, 横尾聡, 宮崎英隆, 対馬義人 : Clinical significance of ^{18}F -alpha-methyl tyrosine (FMT) PET for evaluation of resection area in patients with oral squamous cell carcinoma. 第 72 回日本医学放射線学会総会 4. 13, 横浜

H. 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし

厚生労働科学研究費補助金（医療技術実用化総合研究事業）
分担研究報告書

消化器系癌の臨床試験における主要評価項目（主要エンドポイント）のあり方

研究分担者 竹吉 泉 群馬大学大学院医学系研究科 教授

研究要旨：臨床試験において対象となる癌種、治療法の特徴、デザインの特徴、統計学的な評価法などを十分に吟味した上で、最適なエンドポイントを決定する作業はプロトコル作成上で最も困難な作業の1つである。エンドポイントは、治療の目的に合っており、なおかつ、客観的に評価できる項目が望ましいとされている。臨床試験における治療行為で本来求めたい成果は、死亡率の低下、疾患の発症率の低下、QOLの向上、副作用の低減などであり、これらの評価項目は、主要エンドポイントとなりうる。しかし、それらを治験の期間内で評価することは難しいため、一般には、腫瘍サイズなど短期間で評価できる代用エンドポイントが採用される。FDAの「制がん剤および生物製剤承認のための臨床試験エンドポイント」についての製薬業界向けガイダンスをもとに最適な消化器系癌の臨床試験における主要評価項目について検討し、さらに現在教室で行っている臨床研究で最適なエンドポイントを選んできていたかについて検討した。

消化器系癌の臨床試験における最適な主要エンドポイントを決定することは非常に困難であるが治療法の特徴、デザインの特徴、統計学的な評価法などを十分に吟味した上で、決定しなければならない。その際FDAの製薬業界向けガイダンスが参考となる。

A. 研究目的

臨床試験における評価項目（エンドポイント）とは、「臨床試験の目的に関連する仮説を検証する上で臨床的に意味があり、客観的に評価できる観察・検査項目、またはそれらの合成指標」と定義される（手良向聡、福島雅典：FDA ガイダンス・癌臨床試験のエンドポイント（解説）癌臨床試験のための最適なエンドポイントの選択、臨床評価35（1）128-129, 2007）。エンドポイントは、治療の目的に合っており、なおかつ、客観的に評価できる項目が望ましい

とされている。試験目的に合致した最も重要なものを主要（primary）エンドポイント、それ以外を副次（secondary）エンドポイントと呼ぶ。臨床試験における治療行為で本来求めたい成果は、死亡率の低下、疾患の発症率の低下、QOLの向上、副作用の低減などであり、これらの評価項目は、主要エンドポイントとなりうる（日本薬学会：薬学用語解説）。

しかし、それらを治験の期間内で評価することは難しいため、一般には、腫瘍サイズなど短期間で評価できる代用エンドポイ

ント（サロゲートエンドポイント）が採用される。サロゲートエンドポイントは、治療行為に対する評価を短期間で行うための評価項目である。それ自体では臨床上の利益とならなくても、治療上の成果を合理的に予測しうる場合には、主要エンドポイントとして用いることができる。複数のエンドポイントがある場合には、主要エンドポイントと副次エンドポイントが設定される。主要エンドポイントとは臨床試験において目的とする評価項目であり、薬理的、臨床的に意味のある客観的評価可能な項目が用いられる。副次エンドポイントは、治験の主要な評価項目以外の効果を評価するための項目であり、必ずしも主要エンドポイントとの関連性があるとは限らない（日本薬学会：薬学用語解説）。

2007年5月、「制がん剤および生物製剤承認のための臨床試験エンドポイント」についての製薬業界向けガイダンスがFDAから公布された（日本臨床腫瘍学会：製薬業界向けガイダンス）。FDAは1990年代から大腸癌や肺癌など特定の癌腫における承認のためのエンドポイントについて、米国臨床腫瘍学会や米国血液学会などとワークショップを定期的に共催し、議論を積み重ねて、一般原則に関する集大成としてガイダンスを作成した（小田英世，山中竹春，大門貴志，手良向聡，松井茂之，福島雅典，エンドポイントの臨床的・統計的論点．薬剤疫学．2004；9（suppl）：S28-S29）。また最近では、分子標的薬剤などの新しい作用機序を有する製品の開発が盛んであるが、このガイダンスは臨床的有益性（clinical benefit）の観点から、どのエンドポイントが妥当であるかという議論

を行うための基礎を提供するという意味でも重要とされている（手良向聡，福島雅典：FDAガイダンス・癌臨床試験のエンドポイント（解説）癌臨床試験のための最適なエンドポイントの選択．臨床評価35

（1）128-129，2007）。

対象となる癌種、治療法の特徴、デザインの特徴、統計学的评价法などを十分に吟味した上で、最適なエンドポイントを決定する作業はプロトコル作成上で最も困難な作業の1つである。そこでFDAのガイダンスも鑑みて、主要エンドポイントのあり方について検討したい。さらに、われわれが行ってきた臨床研究および現在行っている臨床研究や先進医療、治験などが最適なエンドポイントを選んできたかについて検討したい。

B. 研究方法

「制がん剤および生物製剤承認のための臨床試験エンドポイント」についての製薬業界向けガイダンスをもとに最適な消化器系癌の臨床試験における主要評価項目について検討し、さらに現在教室で行っている臨床研究や全国規模でおこなっている臨床研究で教室が参加しているもの、そして先進医療について最適なエンドポイントを選んできたかについて検討した。

C. 研究結果

FDAの製薬業界向けガイダンスによれば腫瘍領域におけるエンドポイントは全生存期間、腫瘍評価（無病生存期間（DFS）、客観的奏効率（ORR）、完全奏功、無増悪期間（TTP）、無増悪生存期間（PFS）に基づくエンドポイント、および症状評価に基づくエン

ドポイントがある。そのような観点から我々が行ってきた臨床試験のエンドポイントが最適であったかどうか評価してみる。

以前教室で行った前向き臨床研究として

- ① A Phase I Study of Weekly Paclitaxel and Doxifluridine Combination Chemotherapy in Advanced Gastric Cancer patients (Anticancer Research 25:1291-1296, 2005)
- ② A Phase II Study of Weekly Paclitaxel and Doxifluridine Combination Chemotherapy for Advanced/Recurrent Gastric Cancer (Anticancer Research 31:287-291, 2011)

がある。①に関してはPhase I試験であり、被験薬を少量から段階的に増量し、被験薬の薬物動態や安全性（有害事象、副作用）について検討することを主な目的とした探索的試験とされているため、安全性を評価した本試験のエンドポイントは問題ないものと思われる。②についてはPhase II試験である。Phase II試験は通常第I相の結果をうけて、比較的軽度な少数例の患者を対象に、有効性・安全性・薬物動態などの検討を行う試験とされている。本試験の主要エンドポイントとして奏効率、副次エンドポイントとして全生存率、副作用として行った。全生存率とするより、比較的短時間で結果が出るサロゲートマーカーである奏効率を主要エンドポイントとした。しかしこの場合は副次エンドポイントとした全生存期間を主要エンドポイントとし、奏効率と副作用を副次評価としたほうがよかったと思われる。さらにDFSやPFSについて

検討するようにエンドポイントを組んだほうが良かったように思われる。

現在教室では関連病院と一緒に消化器癌に関する以下の4つの臨床研究を行っている。また先進医療を3件、そして全国規模の大きな臨床試験に3件加わっている。教室が中心となって行っている臨床試験は①Stage II/III大腸癌治癒切除例に対するUFT/LV療法とPSK併用UFT/LV療法の無作為化比較臨床第II相試験 ②切除不能進行再発大腸癌に対する一次治療としてのオキサリプラチン+低用量カペシタビン+ベバズマブ併用療法及びカペシタビン+ベバズマブ維持療法の安全性と有効性の検討：第II相試験③L-OHPを含む初回化学療法に不応・不耐となったKRAS野生型進行再発大腸癌に対する二次治療としてのSIR+Panitumumab併用療法第II相臨床試験④切除不能・再発胃癌に対する2次治療としてのnab-PTX(毎週投与法)の有用性を検討した第II相試験である。

主要エンドポイントは①が3年無病生存割合②③は無増悪生存期間、④は病勢コントロール率としている。副次エンドポイントは①全生存期間、有害事象発生割合、投与完遂割合②全生存期間、全奏効率、奏効期間、安全性、QOL、③奏効率、有害事象、全生存率、治療完遂状況、④奏効率、有害事象、無増悪生存期間、生存期間となっている。

これらの臨床研究に関しては本来最適なエンドポイントは全生存期間であるべきかもしれない。しかし、フォローアップ期間が長期であること、後続のがん治療によって生存解析が交絡を受ける可能性があることなどから、上記のような主要エンドポイント

を組んでいる。

D. 考察

FDA の製薬業界向けガイダンスによれば全生存期間とは無作為化から全原因による死亡までの時間と定義され、intent-to-treat の原則に沿った全患者において測定される。生存期間は最も信頼できるエンドポイントとみなされている。しかし、生存試験を施行、解析する際の困難は、大規模試験においてフォローアップ期間が長期であること、後続のがん治療によって生存解析が交絡を受ける可能性があることなどである。

腫瘍評価によるエンドポイントには、DFS、ORR、TTP、PFS、および治療成功期間(TTF)が含まれる。DFSは無作為化時点から腫瘍再発または全原因による死亡までの期間と定義される。このエンドポイントを最も頻繁に使用しているのは、根治手術または放射線療法後のアジュバント療法においてである。われわれの現在行っている研究ではアジュバント療法については3年無病生存割合を用いている。ORRは、前もって定めた量の腫瘍サイズ縮小を最短期間示した患者の割合と定義される。奏効の持続期間は通常初回奏効時点から文書化された腫瘍増悪までとされる。ORRは薬剤の抗腫瘍活性の直接的尺度であり、単群試験において評価可能である。われわれが以前行った Paclitaxel と Doxifluridine の Phase II Study でも RR を主要評価項目とした。

TTP と PFS は薬剤承認のための主要エンドポイントの役割を果たしている。TTP は無作為化時点から客観的な腫瘍増悪までの時間と定義される。PFS は、無作為化から客観的な腫瘍増悪または死亡までの時間と

定義される。PFS は死亡を含んでいるため、全生存期間とより良く相関しており、患者の死亡は腫瘍増悪と任意に関連するとみなされている。しかし、死亡例の大半ががんと無関係である場合には、TTP は受容できるエンドポイントとされている。PFS は腫瘍増殖を反映することができ、生存利益の判定前に評価できる。その判定は後続治療によって交絡をきたすことがない。治療行為に対する評価を短期間で行うための評価としては非常に有用な方法と思われるし、実際われわれが現在行っている臨床試験でも用いている。しかし、多数の各種既存悪性腫瘍にとって生存期間の代用としての PFS の正式な有効性の立証は困難であり、データは通常、生存期間に対する作用と PFS の相関性の確実な評価を可能にするには不十分とされている。がん試験は小規模な場合が多く、既存薬に証明されている延命効果は通常ささやかなものである。承認を与えるためのエンドポイントとしての PFS の役割は種々のがんの状況によって異なる。PFS 改善が直接的な臨床的有益性を示すのか、それとも臨床的有益性の代用であるのかは、効果の規模と、既存の治療法と比較した新しい治療法のリスク便益の考察に依存すると言われている(日本臨床腫瘍学会:製薬業界向けガイダンス)。

症状改善は臨床的有用性とみなされ、FDA 薬剤承認は患者の症状評価および/または症状改善を表す身体兆候(例、体重増、滲出液の減少)を主要有効性エンドポイントとして使用している。しかし、世界的に用いられている健康に関する QOL(HRQL)の尺度は、抗がん剤承認においては主要有効性エンドポイントとして利用されていない。がん症

状無増悪期間 (Time to progression of cancer symptoms) は TTP と同様のエンドポイントであり、代替エンドポイントとなる可能性があるものというよりむしろ臨床的有用性の直接的尺度であるとされている。症状増悪までの期間を評価するには増悪測定上の諸問題も存在する。がん試験が盲検であることは少ないため、評価はバイアスを被ることが考えられる。腫瘍増悪からがん症状発現までの時間が遅延することもありうる。症状エンドポイントに達する前に代替療法を開始することも多く、この解析に交絡をもたらす。がん試験の多くはがん症状が最小限である患者で行なわれ、腫瘍症状と薬剤毒性を識別することが困難であるとされている (日本臨床腫瘍学会: 製薬業界向けガイダンス)。

消化器系癌については手術、放射線療法、化学療法などの治療法の特徴、デザインの特徴、統計学的な評価法などを十分に吟味した上で、最適なエンドポイントを決定しなければならないが、その際先に述べた FDA の製薬業界向けガイダンスが非常に有益と思われる。

E. 結論

消化器系癌の臨床試験における最適な主要エンドポイントを決定することは非常に困難であるが治療法の特徴、デザインの特徴、統計学的な評価法などを十分に吟味した上で、決定しなければならない。その際 FDA の製薬業界向けガイダンスが参考となる。

F. 健康危険情報

なし

G. 研究発表

戸塚 統, 川手 進, 平井 圭太郎, 小川 博臣, 戸谷 裕之, 吉成 大介, 須納瀬 豊, 竹吉 泉. 経口 UFT+PSK による治療を受けた Stage II/III の大腸癌患者における術前リンパ球比率に関するサブセット解析 癌と化学療法 2013;40:2525-8

五十嵐隆通, 竹吉泉 他. 大腸癌における TS-1+CPT-11 併用療法の使用経験 第 27 回 関越 DIF 研究会 2014. 2 大宮

H. 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

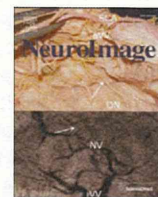
Ⅲ. 研究成果の刊行に関する一覧表

研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Marumo K, Takizawa R, Kinou M, Kawasaki S, Kawakubo Y, Fukuda M, Kasai K	Functional abnormalities in the left ventrolateral prefrontal cortex during a semantic fluency task, and their association with thought disorder in patients with schizophrenia.	<i>NeuroImage</i>	85	518-526	2014
Sato H, Yahata N, Funane T, Takizawa R, Katura T, Atsumori H, Nishimura Y, Kinoshita A, Kiguchi M, Koizumi H, Fukuda M, Kasai K	A NIRS-fMIR investigation of prefrontal cortex activity during a working memory task.	<i>NeuroImage</i>	83	158-173	2013
Takei Y, Suda M, Aoyama Y, Yamaguchi M, Sakurai N, Narita K, Fukuda M, Mikuni M	Temporal lobe and inferior frontal gyrus dysfunction in patients with schizophrenia during face-to-face conversation: a near-infrared spectroscopy study.	<i>J Psychiat Res</i>	47	1581-1589	2013
Kinou M, Takizawa R, Marumo K, Kawasaki S, Kawakubo Y, Fukuda M, Kasai K	Differential spatiotemporal characteristics of the prefrontal hemodynamic response and their association with functional impairment in schizophrenia and major depression.	<i>Schizophr Res</i>	85	498-507	2014
Takizawa R, Fukuda M, Kawasaki S, Kasai K, Mimura M, Pu S, Noda T, Niwa S, Okazaki Y, the Joint Project for Psychiatric Application of Near-Infrared Spectroscopy (JPSY-NIRS) Group	Neuroimaging-aided differential diagnosis of the depressive state.	<i>NeuroImage</i>	85	498-507	2014
Ishijima M, Nakamura T, Shimizu K, Hayashi K, Kikuchi H, Soen S, Omori G, Yamashita T, Uchio Y, Chiba J, Ideno Y, Kubota M, Kurosawa H, Kaneko K	Intra-articular hyaluronic acid injection versus oral non-steroidal anti-inflammatory drug for the treatment of knee osteoarthritis: a multicenter, randomized open-label, non-infertility trial.	<i>Arthritis Research & Therapy</i>	16(1)	in press	2014

IV. 研究成果の刊行物・別冊



Functional abnormalities in the left ventrolateral prefrontal cortex during a semantic fluency task, and their association with thought disorder in patients with schizophrenia

Kohei Marumo ^{a,1}, Ryu Takizawa ^{a,b,*}, Masaru Kinou ^a, Shingo Kawasaki ^{a,c}, Yuki Kawakubo ^d, Masato Fukuda ^e, Kiyoto Kasai ^a

^a Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

^b MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London SE5 8AF, UK

^c Optical Topography Group, Application Development Office, Hitachi Medical Corporation, Tokyo 101-0021, Japan

^d Department of Child Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

^e Department of Psychiatry and Human Behavior, Graduate School of Medicine, Gunma University, Gunma 371-8511, Japan

ARTICLE INFO

Article history:

Accepted 17 April 2013

Available online 24 April 2013

Keywords:

Schizophrenia

Verbal fluency task (VFT)

Near-infrared spectroscopy (NIRS)

Semantic processing abnormality

Prefrontal hemodynamic response

Thought disorder

ABSTRACT

Thought disorder is one of the primary symptoms in schizophrenia, yet the neural correlates and related semantic processing abnormalities remain unclear. We aimed to investigate the relationship between functional prefrontal abnormalities and thought disorder in schizophrenia using 2 types of verbal fluency tasks: the letter fluency task (LFT) and the category fluency task (CFT). Fifty-six adult patients with schizophrenia and 56 healthy controls matched for age, gender, and IQ participated in the study. During completion of the 2 types of verbal fluency tasks, we measured oxy- and deoxy-hemoglobin concentration ([oxy-Hb] and [deoxy-Hb]) signal changes over a wide area of the bilateral prefrontal cortex, using a 52-channel near-infrared spectroscopy (NIRS) system. Thought disorder scores were evaluated using the positive and negative syndrome scale. CFT performance was significantly higher than LFT performance in both groups, while there was no significant difference in any prefrontal NIRS signal changes between the 2 tasks in either group. In both versions of verbal fluency task, healthy controls exhibited a significantly greater NIRS signal change than did patients with schizophrenia. On the CFT only, left ventrolateral prefrontal NIRS [deoxy-Hb] signals were significantly associated with thought disorder scores in patients with schizophrenia. Our results suggest that left ventrolateral prefrontal abnormalities in category fluency might be related to thought disorder in schizophrenia. This could lead to an improved understanding of the neural mechanisms within the left ventrolateral prefrontal cortex involved in mediating semantic processing, as well as the relationship between semantic processing abnormalities and thought disorder in schizophrenia.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Thought disorder, or thought “disorganization,” is considered to be one of the main symptoms of schizophrenia (Bleuler, 1911/1956). It commonly manifests as language that is difficult to make sense of. The severity of thought disorder is heterogeneous in schizophrenia, and cognitive neuropsychological studies suggest that thought disorder may result from semantic processing abnormalities (Aloia et al., 1998; Goldberg et al., 1998). Verbal fluency tasks (VFT) require subjects to produce as many words as possible based on certain cues, and have been widely used to evaluate deficits in semantic processing. Imaging studies (magnetic resonance imaging and positron emission tomography) have

also revealed distinct brain abnormalities related to VFT performance in schizophrenia (Frith et al., 1995; Yurgelun-Todd et al., 1996). One review article on functional neuroimaging (Costafreda et al., 2006) pointed out that VFT consistently elicited activations in the left ventrolateral prefrontal cortex (PFC), which corresponds to Broca's area or the inferior frontal gyrus. While a number of researchers have investigated the relationship between the neural correlates of semantic processing and thought disorder in schizophrenia (Assaf et al., 2006), the details, particularly with regard to VFT, have yet to be elucidated.

The 2 versions of VFT are classified according to the type of cue used: one requires the subject to generate words belonging to a specific semantic category (category fluency task; CFT), while the other requires the generation of words beginning with a specific letter (letter fluency task; LFT). The LFT and CFT make equivalent demands on the executive system (Ruff et al., 1997), with the CFT additionally requiring the involvement of the semantic system (Henry and Crawford, 2005). Recent studies elucidated that CFT performance was one of the candidates for

* Corresponding author at: Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. Fax: +81 3 5800 6894.

E-mail address: takizawa@umin.ac.jp (R. Takizawa).

¹ These authors contributed equally to this work.

the endophenotype of schizophrenia itself (Hurford et al., 2011; Magaud et al., 2010; Snitz et al., 2006). As thought disorder in schizophrenia is strongly associated with impaired semantic processing (Kerns and Berenbaum, 2002; Spitzer, 1997), CFT is thought to be more associated with thought disorder than LFT; however, neuropsychological findings between VFT and thought disorder to date have been somewhat inconsistent (Barrera et al., 2005; Bokar and Goldberg, 2003; DeFreitas et al., 2009; Docherty et al., 2011; Doughty and Done, 2009). Furthermore, no study has investigated the relationship of neural activity during both types of VFT with thought disorder in schizophrenia.

Near-infrared spectroscopy (NIRS) is a non-invasive optical technique, with high temporal resolution (≥ 10 Hz), used to measure the relative concentrations of oxy- and deoxy-hemoglobin ([oxy-Hb] and [deoxy-Hb]) in the cerebral cortex, a measurement reflective of neural activity. This technique allows subjects to be seated comfortably in a well-lit room, and enables continuous, simultaneous measurements to be made over time, thereby allowing the assessment of cortical activity associated with human behavior and cognition. These advantages suggest that NIRS could potentially be developed into a method for bedside monitoring in patients with psychiatric disorders.

To date, the majority of NIRS studies have employed only 1 version of VFT, and have revealed a prefrontal functional abnormality in schizophrenia (Reif et al., 2011; Suto et al., 2004; Takizawa et al., 2008). Three studies (Ehliis et al., 2007; Ikezawa et al., 2009; Kubota et al., 2005) did use both versions of VFT, but evaluated limited areas of the cortex, and did not specifically investigate the relationship between VFT performance and thought disorder in schizophrenia. To expand upon these earlier studies, we covered a wide area of the bilateral prefrontal cortex (PFC) using 52-channel NIRS system in order to investigate the functional abnormalities and their spatiotemporal relationship with thought disorder in schizophrenia.

The objective of the current study was to investigate the relationship between activity in the prefrontal sub-regions during the 2 types of VFTs and thought disorder in patients with schizophrenia, and to test whether NIRS could potentially be used clinically to assess the severity of thought disorder in schizophrenia. We hypothesized that VFT performance and task-related NIRS signals would be attenuated during both versions of the VFT in schizophrenia, and that the prefrontal NIRS signal during the VFT would be associated with thought disorder.

Methods

Subjects

Fifty-six Japanese adult patients with schizophrenia (27 men and 29 women), and 56 healthy Japanese adults (27 men and 29 women), matched for age, participated in the study (Table 1). The exclusion criteria for both groups were neurological illness, traumatic brain injury with any known cognitive consequences or loss of consciousness for more than 5 min, a history of electroconvulsive therapy, and alcohol/substance abuse or addiction. An additional exclusion criterion for the control group was a history of psychiatric disease, or a family history of axis I disorders in any first-degree relatives. Any patients with schizophrenia who had other psychiatric or physical comorbidities were excluded. Patients were recruited from among outpatients and inpatients at the University of Tokyo Hospital. They were diagnosed with schizophrenia according to the DSM-IV criteria (American Psychiatric Association, 1994). All patients were medicated with psychotropic drugs (chlorpromazine equivalent dose 768.3 ± 658.0 mg/day) (see Table 1). None of the patients were in an acute phase, but all had some residual psychiatric symptoms at the time of NIRS measurement. Psychiatric symptoms were evaluated by trained psychiatrists (R.T. and K.K.) using the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987). We calculated the total PANSS scores of the patients by using a five-factor model with confirmatory factor analysis (van der

Gaag et al., 2006a). In this study, the disorganization factor score was used as the primary measure for analysis, as an index of thought disorder. This was the total score of items that showed consistent and unique loadings on the disorganization factor (van der Gaag et al., 2006b).

All subjects were right-handed, as defined by the Edinburgh Inventory (Oldfield, 1971). Premorbid IQs were estimated using the Japanese version of the National Adult Reading Test (Matsuoka et al., 2006). These measures were not significantly different between patient and control groups (Table 1). Socioeconomic status (SES) and parental SES were assessed using the Hollingshead scale (Hollingshead, 1965). Written informed consent was obtained from all subjects prior to their participation in our study, and all protocols were approved by the Ethics Committee of Medicine, The University of Tokyo (Institutional Review Board number: 630).

Activation tasks

Changes in the concentrations of [deoxy-Hb] and [oxy-Hb] were measured during a verbal fluency task, as described previously (Takizawa et al., 2008, 2009a). The task consisted of a 30-s pre-task baseline, a 60-s activation period, and a 70-s post-task baseline. Subjects were instructed to repeat the Japanese syllables /a/, /i/, /u/, /e/, and /o/ during both the pre- and post-task baseline periods. During the LFT, subjects were instructed to produce as many words as possible beginning with a designated syllable, for 20 s each (first: /to/, /a/, or /na/, second: /i/, /ki/, or /se/, and third: /ta/, /o/, or /ha/). In the CFT, subjects were asked to produce as many words as possible within a given semantic category for 20 s each (first: “birds,” “fish,” or “insects”; second: “sweets,” “fruits,” or “vegetables”; third: “vehicles,” “stationery items,” or “home appliances”). Subjects were instructed using an auditory cue at the start and end of the task or baseline period, as well as at the task category change. Before beginning a task session, subjects were given audiovisual, on-screen instructions regarding how to repeat the Japanese syllables during the baseline periods (at a frequency of 1 Hz), and how to provide an answer during the task periods. Each subject performed a practice trial to ensure that they understood the instructions. The total number of correct words produced during the tasks was defined as a measure of task performance. The trial order was counterbalanced among participants.

NIRS measurements

The 52-multi-channel NIRS machine (ETG-4000; Hitachi Medical Corporation) measures relative changes in [oxy-Hb] and [deoxy-Hb] using 2 wavelengths (695 nm and 830 nm) of infrared light, based on the modified Beer–Lambert law. The distance between pairs of detector probes was set at 3.0 cm. A channel (ch) was defined as the measurement area between a pair of source-detector probes. As described previously (Takizawa et al., 2008, 2009b), the probes of the NIRS machine were fixed with 3×11 thermoplastic shells, and placed at 52 measuring points, with the lowest probes positioned along the T4–Fpz–T3 line, according to the international 10–20 system used in electroencephalography. These points are labeled ch 1–52, starting from the right-posterior to the left-anterior (Fig. 2). This arrangement can measure [oxy-Hb] and [deoxy-Hb] from the bilateral prefrontal, and superior temporal cortical surface regions. We estimated the cortical localization of each channel according to the virtual registration method (Tsuzuki et al., 2007; Tzourio-Mazoyer et al., 2002).

Each participant was seated in a chair in front of a desk for all measurements. Subjects were asked to avoid body movements, including head movements, biting, and strong blinking, during the NIRS measurements, in order to minimize artifacts from perfusion changes unrelated to the study task.

The time resolution of the NIRS apparatus was set at 0.1 s [Hb], and changes were analyzed using the first-order correction to exclude changes unrelated to the task, such as very slow oscillations or baseline

Table 1
Demographic characteristics of control subjects and of patients with schizophrenia.

	Healthy controls	Patients with schizophrenia	p-Value
N	56	56	
Age, years	40.9 ± 11.5	40.0 ± 11.0	0.651
Gender, women/men	29/27	29/27	1.000 ^a
Estimated premorbid IQ	104.0 ± 7.6	102.5 ± 11.6	0.490
Handedness	90.2 ± 28.5	93.7 ± 13.1	0.416
Self socio-economic status	1.9 ± 0.7	3.4 ± 1.1	p < 0.001
Parental socio-economic status	2.4 ± 0.8	2.5 ± 0.8	0.468
Age at onset, years		26.3 ± 8.8	
Duration of illness, years		14.1 ± 10.3	
PANSS			
Positive	–	16.7 ± 5.1	
Negative	–	20.9 ± 6.3	
General psychopathology	–	38.5 ± 7.9	
PANSS five factors			
Positive symptoms factor	–	12.3 ± 4.0	
Negative symptoms factor	–	22.1 ± 7.0	
Disorganization factor	–	9.0 ± 2.5	
Excitement factor	–	8.0 ± 2.6	
Emotional distress factor	–	8.9 ± 2.9	
Global assessment of functioning	–	47.4 ± 12.9	
Equivalent dose (mg/day)	–		
Chlorpromazine	–	763.3 ± 658.0	

Abbreviations: IQ, intelligence quotient; PANSS, Positive and Negative Syndrome Scale.

^a Chi-square test was used for testing group difference. Otherwise, *t*-test was used.

drifts. To acquire a stable baseline, a 20-s non-measured period was included in the 30-s pre-task period; NIRS measurement started in the last 10 s of the pre-task period. Pre-task baseline was determined as the mean from the last 10 s of the pre-task period, and the post-task baseline was determined as the mean from the last 5 s of the post-task period; the first order correction was performed between the 2 baselines. The moving average method was applied in order to eliminate short-term motion artifacts (moving average window: 5 s). Data displaying artifacts were excluded separately for each channel using a fully automated procedure as described previously (Takizawa et al., 2008). Briefly, the algorithms that we developed detected channels that showed noise artifacts and body-movement artifacts to ensure the acquisition of more reliable data. We defined three kinds of noise artifacts (high frequency noise, low frequency noise and no signal) and body-movement artifacts show sharp signal changes compared with those of normal hemodynamics (for details see Supplementary material I).

Statistical analyses

In all analyses, a *p*-value < 0.05 was considered to be significant, and an appropriate correction method was applied to multiple comparisons when needed. All values are expressed as mean ± standard deviation (SD). Statistical analysis was performed using SPSS software version 20.0.0J (IBM Corp., 2011, Chicago IL).

Descriptive analyses

Task performance was compared between groups using a Student's *t*-test, and between LFT and CFT scores in each group using a paired *t*-test. Task difference score (Goldberg et al., 1998) was calculated for each participant using the total number of correct words produced for the CFT minus LFT, and then compared between groups using a paired *t*-test.

NIRS analyses

In order to confirm that there was significant activation during the tasks relative to the pre-task baseline, we compared the mean variation of [oxy-Hb] and [deoxy-Hb] during the pre-task and task segment for each channel using a paired *t*-test. Channels with significant activation were termed "effective." Next, for the effective channels, we examined the variation of [oxy-Hb] and [deoxy-Hb] during task segments

between control and patient groups using a Student's *t*-test. To elucidate the spatiotemporal characteristics of functional impairments in schizophrenia, we also calculated the effect size (Cohen, 1988) for each difference in these channels.

When examining several items in a multiple comparison, the significance level becomes more lenient as the number of items being examined increases, and type 1 errors may occur. In order to avoid this, we conducted a *t*-test on effective channels (*m*) and ranked them from the lowest *p* value. The significance level of their rank (*i*) was 0.05 (*i*) / (*m*), and significance was investigated using the False Discovery Rate (FDR) (Verhoeven et al., 2005).

Relationship with thought disorder in schizophrenia

Among the channels showing a significant difference between patients and healthy controls, we sought to investigate the association between the detailed spatiotemporal characteristics of NIRS signal changes, and the disorganization score in schizophrenia. Initially, we investigated the associations between mean NIRS signal change and the disorganization score. If any significant association was found, we confirmed whether these associations were robust using the raw waveform data along the time-course of the NIRS signal. According to the conservative method of Kameyama et al. (2006) and Marumo et al. (2009), in order to avoid multiple comparison errors, those channels which had *p* values of < 0.05 for more than 20 s consecutively (200 comparison time-points in NIRS signals) during the measurement were considered to have a significant correlation.

Results

Descriptive analyses

The demographic characteristics of healthy controls and patients with schizophrenia are shown in Table 1. Age, gender, premorbid IQ, handedness, and parental socio-economic status were not significantly different between the 2 groups.

Descriptive analyses of task performance are summarized in Table 2. For both the LFT and CFT, the number of correct words generated was significantly higher in the control group than in the patient group. Task performance in the LFT was lower than in the CFT in both groups (schizophrenia group; *p* < 0.001, control group; *p* < 0.001), while the task difference score (CFT performance minus LFT performance) of controls was not significantly different from that of patients with schizophrenia.

NIRS wave forms

The grand averaged waveforms of [oxy-Hb] and [deoxy-Hb] during cognitive activations in both healthy controls and patients with schizophrenia are shown in Fig. 1. To confirm significant activations during the task segment, we examined significant changes relative to pre-task baseline in control subjects. For [oxy-Hb], effective channels with significant activations during LFT were ch 12–13, 19, 24, 33–35, 38–42, and 43–52 (FDR *p* < 0.05, corrected with 52 channels), and effective channels in CFT were ch 13, 33, 35, 38–41, 43, 45–46, and 49–52 (FDR *p* < 0.05, corrected with 52 channels) (Fig. 2). For [deoxy-Hb], effective channels in LFT were ch 13, 22–23, 25, 27–31, 34–36, 38–42, 45–46, 49, and 51–52 (FDR *p* < 0.05, corrected with 52 channels), and effective channels in CFT were ch 19–20, 22–24, 27–31, 32–34, 36–42, 43–46, and 49–52 (FDR *p* < 0.05, corrected with 52 channels).

Among these effective channels, we found no significant difference in cortical reactivity (NIRS signal change during the task relative to baseline) between the LFT and CFT in the 2 groups, in any of the channels (FDR *p* > 0.05, n.s.). Significant differences between the 2 groups in terms of [oxy-Hb] and [deoxy-Hb] are shown, together with the effect sizes, in Fig. 3. For [oxy-Hb], channels with significant

differences between the groups were observed within a wide area of the bilateral prefrontal cortex (PFC) during both LFT (ch 12, 19, 24, 34, 35, 38–43, and 45–52) and CFT (33, 35, 38–41, 43, 45, 46, and 49–52) (FDR $p < 0.05$). For [deoxy-Hb], significant differences between the groups were observed during the CFT only in ch 40, 41, 50, and 51 (FDR $p < 0.05$), which were located in the left ventrolateral PFC (Fig. 4), while there was no significant difference during the LFT.

Correlation with thought disorder in schizophrenia

Among the channels showing a significant difference between patients and healthy controls, significant positive correlations between [deoxy-Hb] variations during the CFT and disorganization scores were found in ch 35 ($r = 0.361$), 38 ($r = 0.482$), 41 ($r = 0.326$), 49 ($r = 0.345$), 50 ($r = 0.309$), and 51 ($r = 0.298$), while there was no significant correlation of disorganization scores with [oxy-Hb] variations during the CFT. Additionally, the time-course analysis of the associations during CFT revealed not only that the more severe disorganization symptoms were, the more reduced the [deoxy-Hb] variations were, but also that there were significant associations in the latter half of the time-course of the NIRS measurement in ch 40 ($r = 0.283$ – 0.433), 41 ($r = 0.288$ – 0.403), 50 ($r = 0.273$ – 0.496), and 51 ($r = 0.289$ – 0.352) ($p < 0.05$, with 200 consecutive time-points) (Fig. 4). We found no significant association between thought disorder and either [oxy-Hb] or [deoxy-Hb] variations during LFT.

Discussion

Our results suggest that thought disorder in schizophrenia is associated with left ventrolateral prefrontal dysfunction during semantic processing. To evaluate bilateral PFC activations equally, and to investigate the spatiotemporal characteristics of prefrontal sub-regions, we used 52-channel NIRS, which covers a wide area of the PFC, and a relatively large sample size. Our main findings can be summarized by the following 3 points. (1) There was a difference between control and schizophrenia groups on both types of VFT. However, there was no difference in the NIRS signal changes between the 2 types of VFT for either group. (2) [Oxy-Hb] variation in the PFC, during both the LFT and CFT, was reduced in patients with schizophrenia compared to controls, while we observed reduced [deoxy-Hb] variation in the left ventrolateral PFC during the CFT only in the patient group. There was no such reduction in any PFC regions during the LFT. (3) During the CFT only, the [deoxy-Hb] variation in the same left ventrolateral PFC region was significantly correlated with the disorganization score in schizophrenia. These results suggest that a functional impairment exists in the left ventrolateral PFC in patients with schizophrenia, and that this might be the biological substrate for the semantic processing abnormalities and thought disorder that frequently occurs in schizophrenia.

Performance and NIRS signal during VFT

As expected, in both types of VFT, performance in patients with schizophrenia was significantly lower than in healthy controls, which is consistent with the findings of previous studies (Bokat and Goldberg, 2003; Henry and Crawford, 2005). CFT scores were higher than LFT scores in both the patient and control groups, which is also

consistent with the findings of several previous studies (Elvevag et al., 2001; Kremen et al., 2003), including previous NIRS imaging studies (Ehllis et al., 2007; Ikezawa et al., 2009). However, the only previous NIRS study (Kubota et al., 2005) that employed both types of VFT reported a lack of significant group difference in LFT scores. This might be due to the difference in the task time-course design.

In our study, the task difference score (CFT minus LFT) was not significantly different between the schizophrenia group and healthy controls. This suggests that, in terms of task performance, patients with schizophrenia are not disproportionately deficient in the CFT compared to the LFT. This contradicts a number of previous findings on alphabet-based language speaking in patients with schizophrenia (Bokat and Goldberg, 2003); however, this is often in agreement with the findings in Japanese patients with schizophrenia (Suga et al., 2011; Sumiyoshi et al., 2004), suggesting that the discrepancy could be attributable to differences in semantic processing between speakers of different languages.

From the perspective of NIRS signal change, as expected, we found that in our healthy control subjects, significant activations during both types of VFT occur for both [oxy-Hb] and [deoxy-Hb] signals. To expand the previous NIRS studies that measured changes over a more limited PFC area (Ikezawa et al., 2009; Ehllis et al., 2007; Kubota et al., 2005), we confirmed that these significant activations presented bilaterally across a wide area of the PFC (including ventrolateral, dorsolateral and frontopolar regions) (Fig. 2). At first glance of Fig. 2, NIRS [oxy-Hb] activations during CFT look to be more left-sided, or spread over a smaller area than those during LFT, but in fact, these cortical reactivities in any PFC area were not statistically different between the 2 tasks in either group. This is in contrast with the findings of previous NIRS studies, in which PFC activation during LFT was found to be larger than that during CFT in healthy controls (Ehllis et al., 2007; Kubota et al., 2005). Further investigation is required in order to elucidate the similarities and differences in the neural mechanisms underlying LFT and CFT.

Attenuated NIRS signal in schizophrenia

We confirmed significantly reduced [oxy-Hb] variations in the bilateral PFC during both the LFT and CFT in the schizophrenia group compared to controls, which is in line with the findings of a previous NIRS study (Ehllis et al., 2007). However, 2 other NIRS studies reported different findings; Ikezawa et al. (2009) found no group difference in NIRS [oxy-Hb] signal during CFT, and Kubota et al. (2005) found a more hyperactive activation pattern during CFT than LFT in patients with schizophrenia. This difference might be due to these studies employing 2-channel NIRS that could potentially miss the peak or the trend of PFC activations. As was discussed in our previous publication (Takizawa et al., 2008), the group differences in NIRS [oxy-Hb] signals are widespread across the PFC, including the dorsolateral, ventrolateral and frontopolar areas (Fig. 3), which might reflect impairment of the executive system in schizophrenia (Alvarez and Emory, 2006). Furthermore, the group difference for [oxy-Hb] variations during LFT was more pronounced, and evident across a wider area of the PFC than those during the CFT (see Fig. 3), which is consistent with previous observations (Ehllis et al., 2007). This suggests that the NIRS [oxy-Hb] signal during the LFT may be a more sensitive indicator of this impairment than that during the CFT.

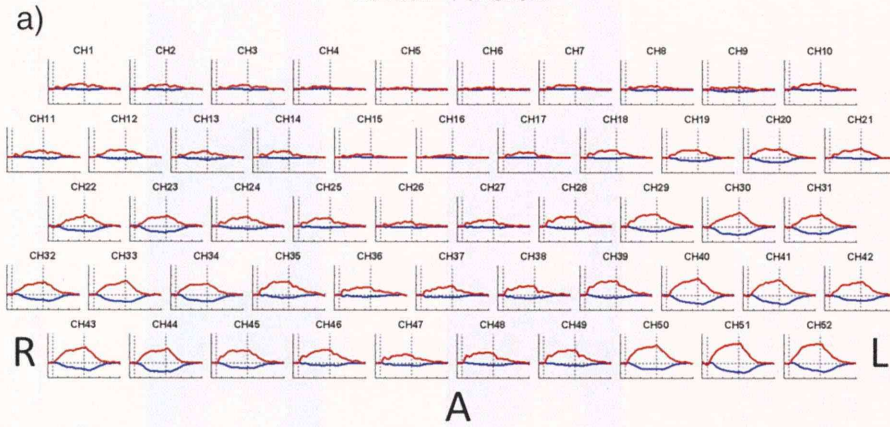
In contrast, we found no group difference in [deoxy-Hb] variations during the LFT in any channel, while a group difference during the CFT was observed in the left ventrolateral PFC (Fig. 4), which is in line with the results reported by Ikezawa et al. (2009). However, 2 studies reported contradictory findings; Ehllis et al. (2007) found a significant group difference, while Kubota et al. (2005) found no group difference in NIRS [deoxy-Hb] signals during either type of VFT. These differences might be due to differing participant characteristics, including sample size. The latter 2 studies included less than 20

Table 2
Summary of task performance.

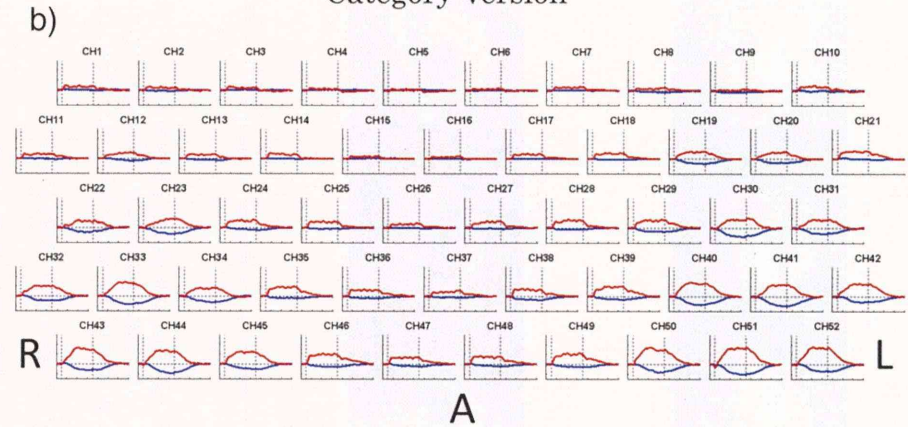
	Healthy controls	Patients with schizophrenia	<i>p</i> -Value
LFT performance	16.6 ± 4.0	14.2 ± 4.5	$p = 0.005$
CFT performance	23.0 ± 4.7	18.8 ± 5.4	$p < 0.001$
Task difference score	6.5 ± 5.5	4.6 ± 5.4	$p = 0.191$

Healthy Controls

Letter Version

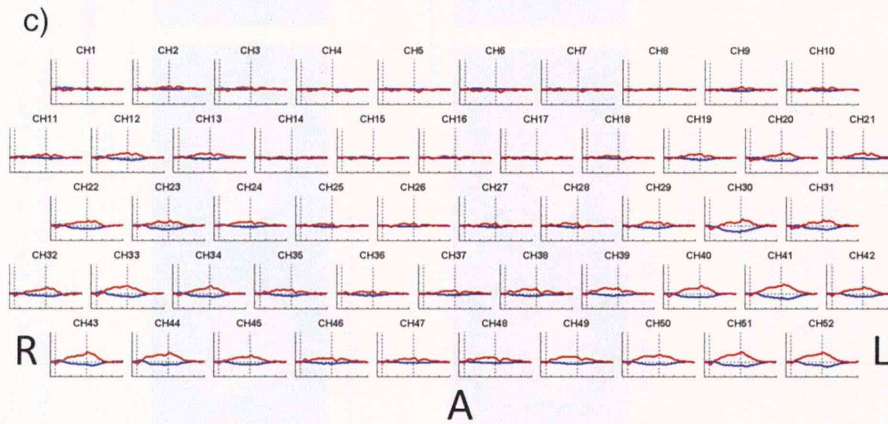


Category Version



Schizophrenia

Letter Version



Category Version

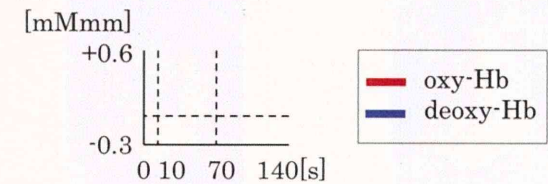
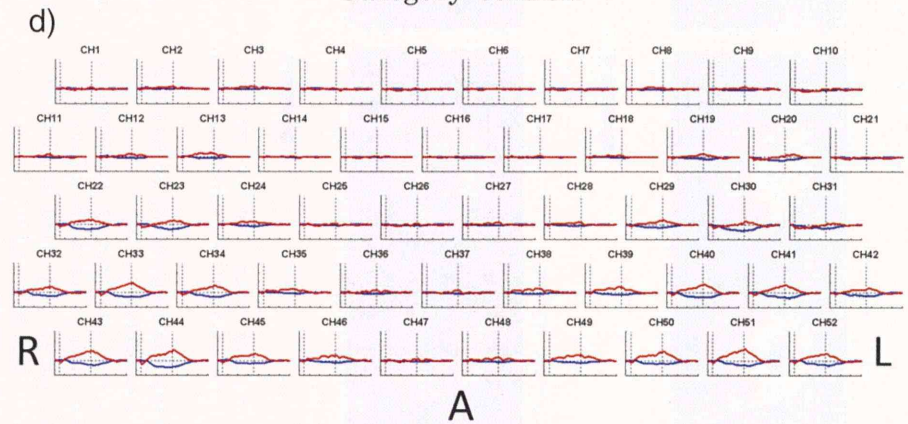


Fig. 1. Grand average waveforms. a. LFT in healthy controls (upper left). b. CFT in healthy controls (upper right). c. LFT in patients with schizophrenia (lower left). d. CFT in patients with schizophrenia (lower right). Oxy- and deoxy-hemoglobin concentration changes during cognitive activation are presented as grand average waveforms in 52 channels in red and blue lines, respectively. Abbreviations; A: anterior; R: right; L: left.

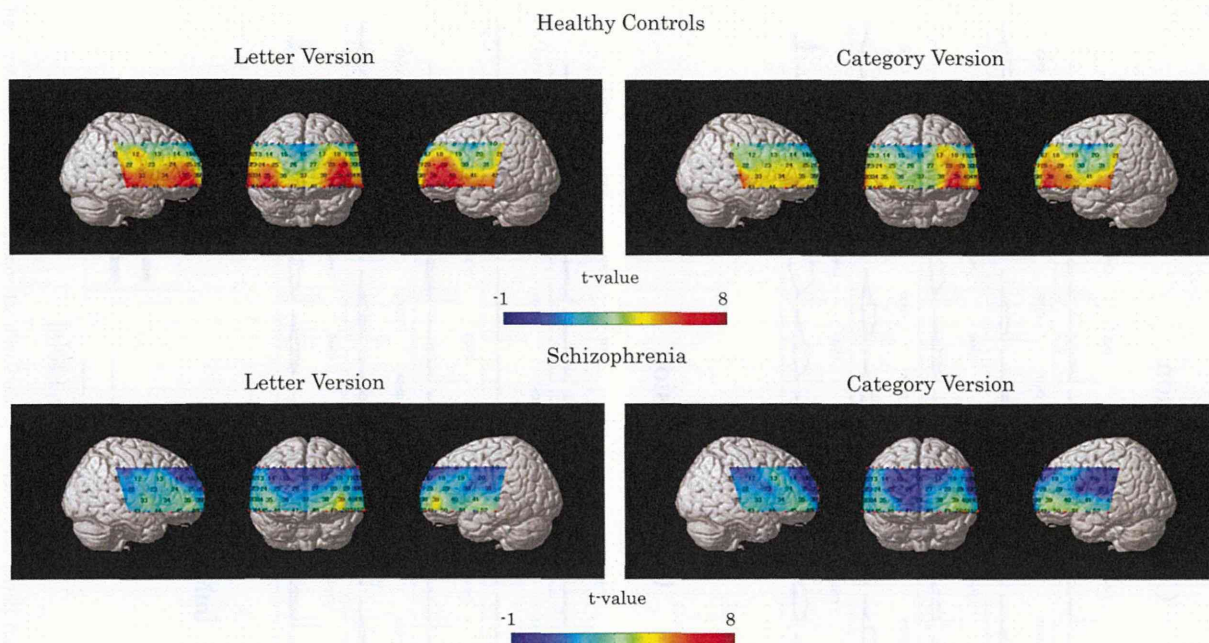


Fig. 2. T-maps for NIRS [oxy-Hb] activations during the task relative to the pre-task baseline for the letter version (LFT; left side) and the category version (CFT; right side) of the verbal fluency task in healthy controls (upper) and in patients with schizophrenia (lower).

subjects per group, while the present study and that by Ikezawa et al. (2009) included more than 30 participants per group.

As suggested by a review article of neuroimaging (Costafreda et al., 2006), the left ventrolateral PFC (Broca's area or the inferior frontal gyrus; Brodmann areas 44, 45, and 47) plays a crucial role during VFT in healthy subjects. The most consistent activations during semantic processing in healthy subjects are known to occur in the left prefrontal cortex (inferior frontal gyrus, anterior cingulate cortex, and dorsolateral prefrontal cortex), in the left temporal cortex, and in the inferior parietal lobule (Cabeza and Nyberg, 2000). Taken together, the finding that this sub-region of the PFC is functionally impaired in schizophrenia is intriguing, and suggests that reduced [deoxy-Hb] variation in the ventrolateral PFC during the CFT may

reflect the semantic processing abnormality in schizophrenia. The NIRS [deoxy-Hb] signal during the CFT in this region may be a more sensitive indicator of this abnormality than that during the LFT.

Association with thought disorder in schizophrenia

In the same left ventrolateral PFC region where the group difference was observed (Fig. 4), we found an association between [deoxy-Hb] variations during the CFT (but not the LFT) and disorganization scores in schizophrenia. As the amplitude of the [deoxy-Hb] signal change was much smaller than that of [oxy-Hb] signal change, cautious attention should be paid to the analyses and interpretations particularly in [deoxy-Hb]. With regard to the correlation analyses, we found significant

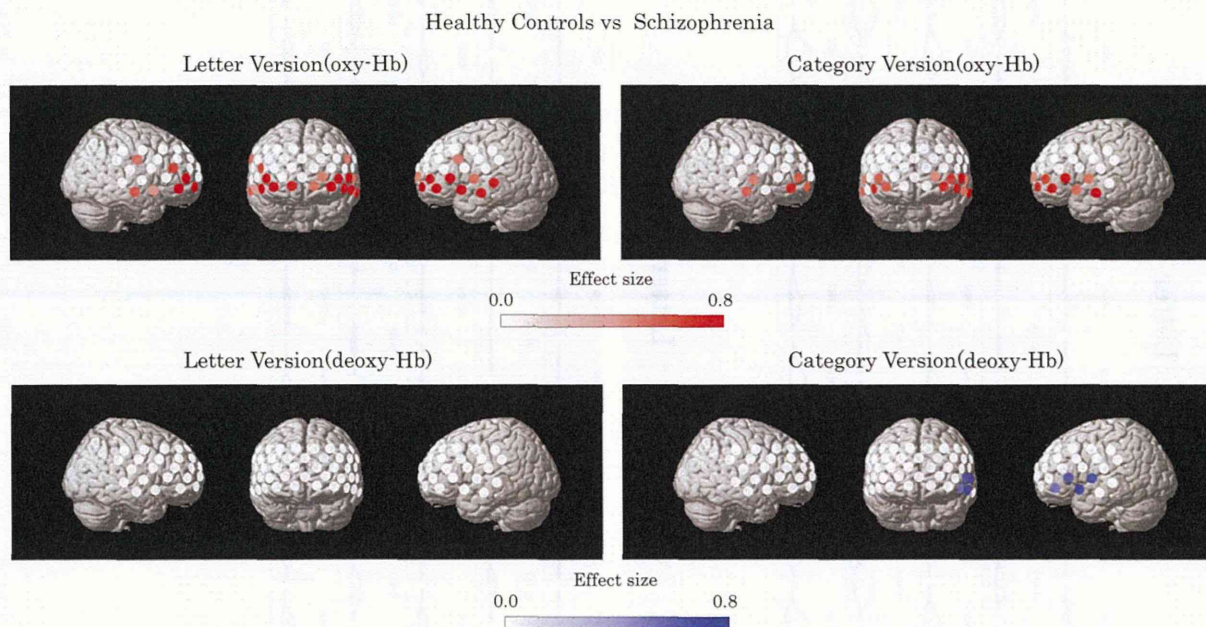


Fig. 3. Group difference in [oxy-Hb] increase (upper) and [deoxy-Hb] decrease (lower) between healthy controls and patients with schizophrenia, elicited by the letter version (left side) and the category version (right side) of the verbal fluency task. Note: Effect sizes of the group difference are indicated by the color gradient. Channels that did not display significant activations are colored in white.

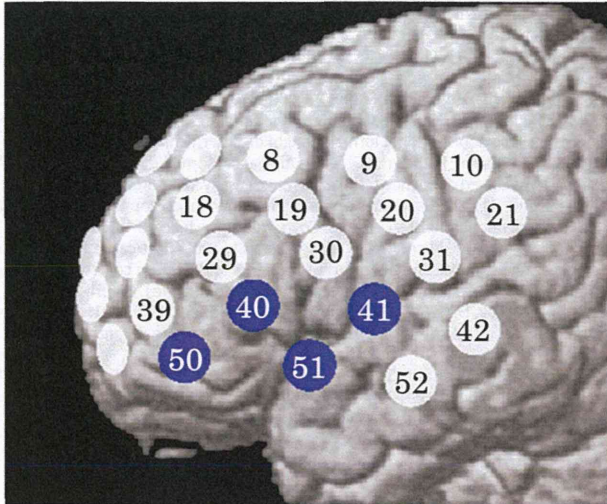


Fig. 4. Measurement positions for the 4 channels are superimposed onto a 3D-reconstructed cortical surface.

correlations in all the prefrontal channels that had significant activations relative to the baseline. Additionally, we sought to confirm these seemingly robust associations by examining the time-course of NIRS signals at each 0.1-s time-point. Even with this time-course analysis, significant correlations in the same direction (in ch 40, 41, 50, and 51) remained in the latter half of the NIRS measurement, in these 4 channels, which are located in the left ventrolateral cortex. Further investigations into time course of NIRS signal should elucidate what it means that the significant correlations were only found in the latter half of the measurement, but the correlation difference between CFT and LFT may be due to the different requirements of each type of VFT.

Although the LFT and CFT make equal demands of the executive system (Ruff et al., 1997), the CFT also requires engagement of the semantic system (Henry and Crawford, 2005). As discussed earlier, the group difference in [deoxy-Hb] observed only during the CFT suggests that the left ventrolateral PFC could be a neural substrate of semantic processing abnormalities in schizophrenia. The finding that the same NIRS [deoxy-Hb] signal in the same left VLPFC region was associated with thought disorder in schizophrenia supports this hypothesis.

Relatively few studies have examined thought disorder in schizophrenia using functional imaging. Despite some inconsistent findings, the overall results indicate that thought disorder is associated with dysfunctions in the temporal, frontal, parietal, and subcortical regions (Kaplan et al., 1993; Kircher et al., 2001; Liddle et al., 1992; McGuire et al., 1998). To the best of our knowledge, only one (Assaf et al., 2006) of these functional imaging studies focused on semantic processing abnormalities in schizophrenia; however, none used 2 types of VFTs. Therefore, this is the first study to demonstrate that the semantic processing abnormality specifically during CFT in left ventrolateral PFC is associated with thought disorganization in schizophrenia.

Limitations

There are several methodological limitations to our study. Firstly, NIRS imaging can provide simultaneous measurements of concentration changes for 2 types of hemoglobin, while the blood oxygenation level-dependent (BOLD) signal measured by functional magnetic resonance imaging (fMRI) reflects only changes in deoxy-hemoglobin. Some authors have suggested that the NIRS [deoxy-Hb] signal is the most appropriate parameter of brain activity (Obrig and Villringer, 2003), because hemoglobin concentration changes depending on oxygen consumption, with brain activity being reflected by decreased [deoxy-Hb], caused by the inflow of newly oxygenated blood. However, based on the results of animal work (Hoshi et al., 2001) and some

studies that have provided evidence of a stronger correlation between the NIRS [oxy-Hb] signal and the BOLD signal measured by fMRI (Strangman et al., 2002), some uncertainty remains regarding the precise relationship between the fMRI BOLD signal and the NIRS [deoxy-Hb] signal. Therefore, further work, particularly regarding [deoxy-Hb] changes, needs to be performed. Secondly, the potential effects of task performance and medication could not be completely excluded, even though we confirmed that these factors were not associated with any NIRS signal changes, consistent with previous NIRS studies (Ehlis et al., 2007; Takizawa et al., 2008). In addition, we confirmed that the significant findings were unchanged if either task performance or medication was included as a covariate in the analysis. Thirdly, although the spatial resolution of NIRS measurement is relatively low, the measurement position might have varied slightly among the participants. We confirmed that there was no significant group difference in T3–T4 distance using the international 10–20 system (healthy controls: 30.2 ± 1.5 cm, patient group: 30.5 ± 1.6 cm; $p > 0.05$, n.s.). Additionally, for the purpose of estimating the cortical localization of each channel, use of the virtual registration method (Tszuki et al., 2007; Tzourio-Mazoyer et al., 2002) enabled the probabilistic registration of NIRS data onto the Montreal Neurological Institute coordinate-space, without measurement of probe positions or MRIs. Using to this method, we confirmed that the 4 channels of particular interest (ch 40, 41, 50, and 51) were approximately located in left inferior frontal gyrus (ventrolateral PFC) (cf., Fig. 3). Two channels have more than half probability estimate to be in the left inferior frontal gyrus (ch 40, 100%; ch 50, 64%), while ch 41 and ch 51 have a greater probability estimate in the left superior temporal gyrus (ch 41, 46%; ch 51, 74%). According to evidence from neuroimaging studies (Cabeza and Nyberg, 2000), it would not be paradoxical if the semantic processing abnormality in schizophrenia was found to also involve the superior temporal gyrus (Shenton et al., 1992).

Implications for future research and clinical practice

As the severity of thought disorder is heterogeneous in schizophrenia, using the NIRS signal in the left ventrolateral PFC as an objective heterogeneity index in schizophrenia might prove useful in elucidating striking effects in genetic or pharmacological studies. In clinical practice, as NIRS is a non-invasive and portable device, and can be easily used to make repetitive measurements, clinicians might objectively evaluate the severity of thought disorder using NIRS [deoxy-Hb] signals in the left ventrolateral PFC, and could monitor it through the course of treatment, provided our findings proved to be replicable in a longitudinal study.

Conclusions

In this study, we have investigated the relationship between PFC function and thought disorder in patients with schizophrenia by using 2 types of verbal fluency tasks and 52-channel NIRS imaging. We found reduced activation in the left ventrolateral PFC in patients with schizophrenia, and a correlation between thought disorder score and [deoxy-Hb] variation in this region. These results suggest that a functional impairment exists in the left ventrolateral PFC in patients with schizophrenia, and that such an impairment may be the biological basis for semantic processing abnormalities and thought disorder.

Funding

This work was supported in part by grants from Grant-in-Aid for Scientific Research (Innovative areas Nos. 23118001 and 23118004 [Adolescent Mind & Self-Regulation] to KK; No. 23791309 to RT) and from the “Development of Biomarker Candidates for Social Behavior” study carried out under the Strategic Research Program for Brain Sciences (to KK) by the MEXT. This study was also supported in part by