

一般口演 | 医療支援

一般口演2 医療支援

2021年11月19日(金) 09:10 ~ 11:10 F会場 (2号館2階224)

[2-F-1-02] AIを用いた病院流動食の摂取量推定システムの臨床現場における精度検証

*田木 真和¹、田尻 真梨²、濱田 康弘³、若田 好史⁴、単 暁⁴、尾崎 和美⁵、久保田 雅則⁶、天野 宗佑⁶、廣瀬 隼¹（1. 徳島大学大学院医歯薬学研究部医療情報学分野, 2. 徳島大学病院 栄養部, 3. 徳島大学大学院医歯薬学研究部疾患治療栄養学分野, 4. 徳島大学病院 病院情報センター, 5. 徳島大学大学院医歯薬学研究部口腔保健支援学分野, 6. foo.log株式会社）

*Masato Tagi¹, Mari Tajiri², Yasuhiro Hamada³, Yoshifumi Wakata⁴, Xiao Shan⁴, Kazumi Ozaki⁵, Masanori Kubota⁶, Sosuke Amano⁶, Jun Hirose¹（1. 徳島大学大学院医歯薬学研究部医療情報学分野, 2. 徳島大学病院 栄養部, 3. 徳島大学大学院医歯薬学研究部疾患治療栄養学分野, 4. 徳島大学病院 病院情報センター, 5. 徳島大学大学院医歯薬学研究部口腔保健支援学分野, 6. foo.log株式会社）

キーワード：AI, Convolutional Neural Network, food intake, nutrition management, liquid food

栄養不良は入院の長期化や合併症の発生リスク増加につながるため、患者の栄養状態を正確に評価し栄養不良を防ぐ必要がある。臨床現場では、簡便性を重視して主に目測法が食事摂取量評価に用いられているが正確性に課題がある。そこで我々は簡便性を維持しつつ、より正確な測定を目指して、食事摂取量を推定するAIモデルを開発中である。料理画像の目測推定と比較したこれまでの検証でAIモデルの有用性が示されたので、本研究では直接料理を見て評価する目測推定と比較した場合の精度を検証した。2020年11月から2021年3月まで徳島大学病院の2つの病棟において、入院患者に提供された普通流動食100食を評価対象とした。それぞれの料理に対して、看護師が直接料理を見て食事摂取量を推定する目測推定と、タブレット端末で撮影した食後の料理画像からAIが食事摂取量を推定するAI推定による評価を行い、食事前後の重量から測定した実際の摂取量（秤量測定）に対する平均絶対誤差、決定係数 R^2 及び正解率を用いて推定精度を比較した。AI推定の平均絶対誤差は0.91で、目測推定の0.17より大きかった。AI推定の決定係数 R^2 0.62と正解率44.7%は、目測推定のそれぞれ0.95, 88.7%よりともに低かった。臨床現場で料理を直接見て評価した結果、先行研究の料理画像評価とは逆に、AI推定の精度は目測推定と比較して低くなった。その理由として、秤量測定による実測値0に対するAI推定の誤差が大ききことと、実測値10に対する目測推定の正解率がほぼ100%に対しAI推定は51%で大きな差があることが挙げられる。しかしながら、AI推定の秤量測定との誤差は臨床現場で望ましいとされている1.00以内だったため、判明した課題である実測値0と10を確実に推定できるようAIモデルを改良することにより精度向上が期待でき、現在その取組みを進めている。

AIを用いた病院流動食の摂取量推定システムの臨床現場における精度検証

田木 真和^{*1}、田尻 真梨^{*2}、濱田 康弘^{*3}、若田 好史^{*4}、単 暁^{*4}、尾崎 和美^{*5}、久保田 雅則^{*6}、
天野 宗佑^{*6}、廣瀬 隼^{*1}

*1 徳島大学大学院医歯薬学研究部医療情報学分野、*2 徳島大学病院 栄養部、
*3 徳島大学大学院医歯薬学研究部疾患治療栄養学分野、*4 徳島大学病院 病院情報センター、
*5 徳島大学大学院医歯薬学研究部口腔保健支援学分野、*6foo.log 株式会社

Clinical validation of an AI-based system for estimating intake hospital liquid food

Masato Tagi^{*1}, Mari Tajiri^{*2}, Yasuhiro Hamada^{*3}, Yoshifumi Wakata^{*4}, Xiao Shan^{*4}, Kazumi Ozaki^{*5}, Masanori Kubota^{*6}, Sosuke Amano^{*6}, Jun Hirose^{*1}

*1 Department of Medical Informatics, Tokushima University Graduate School of Biomedical Sciences, *2 Division of Nutrition, Tokushima University Hospital, *3 Department of Therapeutic Nutrition, Tokushima University Graduate School of Biomedical Sciences, *4 Medical IT Center, Tokushima University Hospital, *5 Department of Oral Health Care Promotion, Tokushima University Graduate School of Biomedical Sciences, *6 foo.log Inc

Accurate assessment of the nutritional status is important for hospitalized patients who are malnourished with higher risk for some complications. Visual estimation method is popular to measure food intake in a clinical environment, however it is not enough reliable from the point of view of accurate measurement. Therefore, we are developing an AI-based system to estimate food intake for more accurate measurement. The purpose of this study is to validate the accuracy of the system compared to visual estimation performed in a clinical environment, since previous validations compared to visual estimation of food images have shown the usefulness of the AI model. We compare the accuracy of the evaluation by AI-based system (AI estimation), the evaluation by looking at food images (image-visual estimation) and the evaluation by directly looking at food (direct-visual estimation) using accuracy score, the mean absolute error and coefficient of determination R^2 at Tokushima University Hospital. The result shows that the mean absolute errors of AI estimation were large and accuracy score and the coefficients of determination R^2 were small, which indicates that the accuracy of AI estimation was low, contrary to the results of previous studies. Further improving would be required to use an AI-based system for estimating intake hospital liquid food in a clinical environment.

Keywords: AI, Convolutional Neural Network, food intake, nutrition management, liquid food

1. 背景

入院患者の 20%から 50%は栄養不良であり、高齢者やがん治療患者ではその割合がさらに高い¹⁾。そして栄養不良は、入院期間の長期化、褥瘡や感染症などの合併症の発症リスクの増加の要因となる²⁾。よって栄養不良を防ぐことが重要で、そのためには患者の栄養状態を評価する必要がある³⁾。

栄養評価の基準の一つに患者の食事摂取量があり、食事摂取量を把握することで栄養介入による改善が期待できる。特に流動食を摂取している患者は、術後や治療による絶食後の状態であり、正確に食事摂取量を把握しなければならない。食事摂取量を測定するために臨床現場では、簡便性を重視して主に目測法を用いているが正確性に課題があることが指摘されている⁴⁾。そこで我々は、簡便性を維持しつつより正確な測定を目指して、Convolutional Neural Network を活用した食事残量を推定する AI モデルを開発し、管理栄養士による料理画像の目測推定と比較した検証でその有用性を示した⁵⁾。しかし、AI システムに関する先行研究では、研究開発環境から臨床現場での実用に移行する際にいくつかの障害が発生するといわれている⁶⁾。実際に臨床現場で AI システムを評価したとき、報告されていた予測精度から大きく低下していたという報告があり⁷⁾、我々の開発した AI モデルについても臨床現場における精度を検証する必要があると考えられる。

2. 目的

本研究では、先行研究で開発した AI モデルを用いた食事摂取量推定システムを構築し、病棟で実際に患者が食事した流動食の評価に利用することで、臨床現場における AI の精度を検証した。

3. 方法

徳島大学病院の 2 つの病棟において、2020 年 11 月から 2021 年 3 月の期間で入院患者に提供された普通流動食(図 1)100 食を評価対象とした。それぞれの料理に対して、食後に写真撮影し AI モデルによる料理写真の画像解析評価(AI 推定)、同じ料理写真を管理栄養士が見て行う目測評価(画像目測推定)、看護師が直接料理を見て行う目測評価(直接目測推定)を実施し、それぞれ食事摂取量を 0 から 10 の 11 段階で推定した値を用いた。それらの値と食事前後の食べ物の重量の差分をデジタルスケールで測る秤量法で実測した値と比較することでそれぞれの推定精度を測定した。秤量法で実測した摂取量は、摂取量が 11 段階となるように換算し、各料理の摂取状態を全く食べていない(摂取無)、一部食べ残しがある(摂取一部)、食べ残しがない(摂取全量)に分類した(表 1)。



図 1 入院患者に提供された普通流動食

表 1 摂取分類と秤量法による食事摂取量の実測換算値

分類	換算値	食事摂取量
摂取無	0	全量の 5%未満の摂取量
摂取一部	1	全量の 5%以上 15%未満の摂取量
	2	全量の 15%以上 25%未満の摂取量
	3	全量の 25%以上 35%未満の摂取量
	4	全量の 35%以上 45%未満の摂取量
	5	全量の 45%以上 55%未満の摂取量
	6	全量の 55%以上 65%未満の摂取量
	7	全量の 65%以上 75%未満の摂取量
	8	全量の 75%以上 85%未満の摂取量
	9	全量の 85%以上 95%未満の摂取量
摂取全量	10	全量の 95%以上の摂取量

3.1 データセット

普通流動食の献立は主食、副食 2 種類、パック飲料、調味料の組み合わせで構成されており、料理の内訳と AI 学習に用いた料理画像の枚数を表 2 に示す。学習に用いた料理画像は、各料理の摂取量が 0 割から 10 割をすべて含むように作成した分量の組み合わせで構成されている。アノテーションツールを利用し、トレイ上に配膳された複数の料理の領域、名称、残量のラベル付けを行い、料理毎に画像を分割した。

表 2 料理名と学習画像枚数

種別	料理名	学習枚数
主食	重湯	504
	清汁	144
	野菜スープ	360
	味噌スープ	144
	赤だし	66
	カルピス	72
	ピーチ果汁	72
	グレープ果汁	72
	オレンジ果汁	72
	ミックスジュース	66
副食	フルーツミックス	66
	牛乳	504
	幼児牛乳	66
	幼児リンゴ	66
	幼児オレンジ	66
パック飲料	無添加充実野菜	66
	塩	504

3.2 食事摂取量推定システム

食事摂取量推定システムの概要を図 2 に示す。本システムは、管理者が献立を事前に登録することで、写真撮影された料理と献立の料理名を一致させる。利用者は、食後の料理をタブレット端末で写真撮影し、その画像データを食事摂取量推定システムサーバへアップロードする。アップロードされた画像データは、AI 画像解析で料理の位置、名称、残量が解析され、判定された内容が画面表示される。タブレット端末からサーバへのデータ送信はインターネットを介して行い、暗号化は Transport Layer Security (TLS) を用いた。



図 2 食事摂取量推定システムの概要

3.3 精度評価

AI 推定、画像目測推定、直接目測推定のそれぞれの評価法による推定値を秤量法による実測値と比較した。精度評価の評価指標には正解率、平均絶対誤差、決定係数 R^2 を用いた。決定係数 R^2 は、全てを実測値の平均値で推定した場合と比べてどの程度誤差が小さいか表す指標で、1 に近づくほど精度が高いことを示している。さらに、正解率と平均絶対誤差を用いて料理の摂取分類別の精度比較を行った。また、絶対誤差の分布を評価するため、推定値と実測値の混合行列表を作成した。

4. 結果

画像目測推定は徳島大学病院の管理栄養士 6 名が実施し、直接目測推定は徳島大学病院の看護師 9 名が実施した。評価した料理は、主食の重湯と副食の野菜スープ、カルピス、ピーチ果汁、グレープ果汁、オレンジ果汁、ミックスジュースで食事摂取量別の料理数を表 3 に示す。

表 3 食事摂取量別の料理件数

分類	摂取量換算値	料理数 n = 300
摂取無	0	23
	1	6
	2	3
	3	2
	4	9
摂取一部	5	6
	6	6
	7	6
	8	3
	9	6
摂取全量	10	230

AI 推定の正解率と決定係数は、画像目測推定と直接目測推定のそれぞれの値より小さかった(表 4)。さらに、AI 推定の平均絶対誤差は 0.91 で、画像目測推定の 0.44 と直接目測推定の 0.17 より大きかった。また、画像目測推定の正解率と決定係数は直接目測推定のそれぞれの値より小さく、平均絶対誤差も大きかった。ただし、平均絶対誤差はどの評価法もすべて 1.0 未満で小さい値を示した。

表 4 評価法別の精度評価

評価法	正解率 (%)	平均 絶対誤差	決定係数 R ²
AI 推定	44.7	0.91	0.62
画像目測推定	82.0	0.44	0.80
直接目測推定	88.7	0.17	0.95

摂取分類別に分けた精度評価の結果を表 5 に示す。摂取無では、直接目測推定の正解率は 95.7%と高かったが、AI 推定と画像目測推定の正解率は低く、平均絶対誤差も直接目測推定のみ小さい値だった。摂取一部では、正解率は各評価とも 30%前後と低く、平均絶対誤差は画像目測推定が最も大きい値を示した。摂取全量の直接目測推定と画像目測推定の正解率はほぼ 100%で高かったが、AI 推定のみ 51.3%と低かった。摂取全量の平均絶対誤差は、どの評価法も 1.0 未満で小さい値を示した。また、評価した料理件数の 77%が摂取全量に分類され、患者の食事摂取量に偏りがあった。

表 5 摂取分類・評価法別の精度評価

分類	評価法	正解率 (%)	平均 絶対誤差
摂取無 n=23	AI 推定	13.0	2.91
	画像目測推定	34.8	1.74
	直接目測推定	95.7	0.39
摂取一部 n=57	AI 推定	27.7	1.19
	画像目測推定	27.7	1.80
	直接目測推定	31.9	0.87
摂取全量 n=230	AI 推定	51.3	0.66
	画像目測推定	97.8	0.03
	直接目測推定	99.6	0.004

推定値と実測値の混合行列表をみると、AI 推定では実測値が 10 の時に誤差の分布のばらつきが多くみられ、誤差のある評価件数が多かった(図 3)。画像目測推定でも全体的にばらつきがみられ、実測よりも摂取量が多いと推定する評価が多かった。一方、直接目測推定では誤差が小さく実測値とほぼ一致していた。特に実測値が 0 の時、直接目測推定では誤差のある評価がほとんどみられなかったが、AI 推定と画像目測推定では誤差の分布のばらつきが多くみられた。

AI 推定

		estimated value										
measured value		0	1	2	3	4	5	6	7	8	9	10
	0	3	2	4	4	5	5	0	0	0	0	0
	1	0	0	4	2	0	0	0	0	0	0	0
	2	0	0	0	0	1	2	0	0	0	0	0
	3	0	0	0	0	1	1	0	0	0	0	0
	4	0	0	0	1	4	1	2	1	0	0	0
	5	0	0	0	0	0	1	2	2	1	0	0
	6	0	0	0	0	2	0	1	1	2	0	0
	7	0	0	0	0	0	0	1	1	3	1	0
	8	0	0	0	0	0	0	1	0	1	1	0
	9	0	0	0	0	0	0	0	0	1	5	0
	10	0	0	0	0	0	0	1	9	18	84	118

画像目測推定

		estimated value										
measured value		0	1	2	3	4	5	6	7	8	9	10
	0	8	6	3	2	2	1	0	0	0	1	0
	1	2	2	0	0	0	1	1	0	0	0	0
	2	1	0	0	0	0	1	0	1	0	0	0
	3	0	0	0	0	1	0	1	0	0	0	0
	4	0	0	1	0	3	1	0	2	1	1	0
	5	0	0	0	0	0	1	1	0	2	1	1
	6	0	0	0	0	0	1	2	0	3	0	0
	7	0	0	0	1	0	0	0	0	1	2	2
	8	0	0	0	0	0	0	0	0	2	1	0
	9	0	0	0	0	0	0	0	0	0	3	3
	10	0	0	0	0	0	0	0	0	2	3	225

直接目測推定

		estimated value										
measured value		0	1	2	3	4	5	6	7	8	9	10
	0	22	0	0	0	0	0	0	0	0	1	0
	1	5	0	1	0	0	0	0	0	0	0	0
	2	0	1	2	0	0	0	0	0	0	0	0
	3	0	0	1	0	0	0	0	1	0	0	0
	4	0	1	0	1	3	3	1	0	0	0	0
	5	0	0	0	0	1	2	1	2	0	0	0
	6	0	0	0	0	0	1	2	3	0	0	0
	7	0	0	0	0	0	0	2	1	2	1	0
	8	0	0	0	0	0	0	0	1	2	0	0
	9	0	0	0	0	0	0	0	0	0	3	3
	10	0	0	0	0	0	0	0	0	0	1	229

図 3 推定値と実測値の混合行列表

5. 考察

臨床現場で入院患者に提供された流動食を評価した結果、AI 推定の平均絶対誤差は 0.91 で、臨床現場で望ましいとされている秤量法との誤差 1.0 以内⁸⁾の指標より小さい値となり、摂取一部の場合に AI 推定の方が画像目測推定よりも誤差は小さくなった。しかしながら、先行研究の結果⁹⁾とは逆に、摂取全量および摂取無では AI 推定の精度は直接目測推定と比較して低く、画像目測推定よりも低かった。AI 推定と目測推定の精度に差ができた理由として、摂取全量に対する各目測推定の正解率がほぼ 100%に対し AI 推定は半分の 51.3%で低かったことが挙げられる。評価した料理の 77%が摂取全量に分類されたため、この正解率の差が精度の差につながったと考えられる。AI 推定の正解率が低かったのは、摂取量 10 の学習用に用いた料理画像に原因があった可能性がある。摂取量 10 は一度全量入った料理を患者がすべて食べて無くなった状態だが、今回学習用に準備した料理画像では料理を全く入れていない食器を写真撮影していた。一度料理が入った状態から食べて無くなった場合、食器内側の側面や底面に少し料理が付いた状態となる。料理を全く入れていない食器ではその状態にならなかったため、特に摂取量 10 から 7 を区別するための画像特徴を学習時にうまく発見できず、誤差が発生したと考えられる。

摂取分類によってそれぞれの精度評価が大きく異なる結果(表 5)となったことから、料理の摂取量によって摂取量推定難度が変わると推測される。摂取一部の正解率は、どの評価法においても 30%前後となったため、推定難度が高い。このように難度が高い料理の摂取量を推定する場合には先行研究の結果と同じように画像目測推定よりも AI 推定の方が誤差は小さくなったが、直接目測推定よりは誤差が大きかった。また、摂取無に対する AI 推定の正解率が 13%、画像目測推定の正解率が 34.8%で、誤差の分布にも大きなばらつきがみられたので(図 3)、この分類も問題の難度が高いと考えられたが、直接目測推定の正解率はほぼ 100%だった。そのため、直接料理をみた場合には写真では捉えられない特徴など他の情報も加えて食事摂取量を判断することで難度が下がった可能性が示唆される。AI 推定の平均絶対誤差は摂取無の 2.91 と一部摂取の 1.19 で差がみられたが、画像目測推定は摂取無の 1.74 と一部摂取の 1.80 でほぼ同等だった。このことから、AI 推定は摂取無で誤差が大きくなることが課題として残る。

本研究の限界として、AI 推定に写真撮影した料理画像を用いた点が挙げられる。料理実物を直接目測推定の方が料理画像を目測推定するより秤量法との相関が高いことが報告されており⁹⁾、本研究も同じように直接目測推定の方が料理画像を用いた AI 推定や画像目測推定より精度が高くなった。また、評価した料理の 77%が摂取全量に分類され、評価件数の分類に偏りがあった。そのため、摂取無と摂取一部の評価した料理の件数が少なくなり、評価に十分な件数でなかった可能性がある。今回は、単一施設での評価に限定して行った。患者に提供される病院流動食のメニューや食器は施設によって異なり、施設ごとの精度や学習画像枚数などを調査する必要がある。また、AI を用いた測定方法の利便性についての評価がまだできていない。精度が高くても使い勝手の悪いシステムでは、臨床現場での実用に移行する際に障害となる⁶⁾。日常的な栄養管理について、AI 画像解析を用いて食事記録を支援するシステムは、簡便さなどの有用性が高いと評価されている¹⁰⁾。臨床現場においても、AI を用いた測定が医療スタッフの利便性を向上させるか評価する必要がある。

6. 結語

AI を用いた食事摂取量推定システムで患者の食事摂取量を推定した結果、秤量法との誤差は臨床応用の許容範囲内だったが、直接目測推定および画像目測推定と比較し精度が低かった。目測推定との精度差の課題が明確になったことから、臨床現場で利用するためには、さらに精度を向上させる必要がある。今後、課題となっている精度改善を図り、より正確かつ簡便に病院流動食の食事摂取量を測定できる AI システムの構築を目指す。

参考文献

- 1) Norman K, Pichard C, Lochs H, et al. Prognostic impact of disease-related malnutrition. Clin Nutr 2008 ; 27(1) : 5-15.
- 2) Correia M I, Heazi R A, Higashiguchi T, et al. Evidence-based recommendations for addressing malnutrition in health care: an updated strategy from the feedM.E. Global Study Group. J Am Med Dir Assoc 2014 ; 15(8) : 544-50.
- 3) Kondrup J, Allison S P, Elia M, et al. ESPEN Guidelines for Nutrition Screening 2002. Clin Nutr 2003 ; 22(4) : 415-21.
- 4) Castellanos V H, Andrews Y N. Inherent flaws in a method of estimating meal intake commonly used in long-term-care facilities, J Am Diet Assoc 2002 ; 102(6) : 826-30.
- 5) 田本真和, 田尻真梨, 濱田康弘ら. AI を用いた病院流動食の残量推定の検証. 日本医療情報学会春季学術大会 2021 ; 25 : 70-1.
- 6) Beede E, Baylor E, Hersch F, et al. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. CHI 2020 : 1-12.
- 7) Wong A, Otles E, Donnelly J P, et al. External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients. JAMA Intern Med 2021 ; 181(8) : 1065-70.
- 8) Sharma M, Rao M, Jacob S, et al. Validation of 24-hour dietary recall a study in hemodialysis patients. J Ren Nutr 1998 ; 8(4) : 199-202.
- 9) Williamson D A, Allen H, Martin P D, et al. Comparison of digital photography to weighed and visual estimation of portion sizes, J Am Diet Assoc 2003 : 103(9) : 1139-45.
- 10) Aizawa K, Maeda K, Ogawa M, et al. Comparative Study of the Routine Daily Usability of FoodLog: A Smartphone-based Food Recording Tool Assisted by Image Retrieval, J Diabetes Sci Technol 2014 ; 8(2) : 203-8.