

## NDB 粗集計 (NDB コードブック) につきまして

本コードブックは、2016 年後半時点に取得可能であった医科 NDB の全項目について、どのような値がセットされているのかを、要約統計量と項目値サンプルを用いて示したものです (特定健診及び歯科は含まれません)。

本コードブックは、2016 年に当講座が NDB の医科全体を取り扱おうと試みた際に、NDB の各変数にどのような値が含まれるのかの一覧を手元に置くことをめざして作成されました。

本コードブックは、「各変数の要約統計量」と「各変数の項目値上位 50 サンプル」の 2 つから構成されています。

たとえば、「表 2 粗集計 医科 (入院) RE」は、NDB の医科入院レセプトの RE レコードに含まれるすべての変数を対象に、要約統計量を示したものです。項番 6「予備 1 (年齢)」は最小値が 1、最大値が 19 で、欠損値割合が 0% ですので、NDB の医科入院レセプトにおける年齢は、実際の年齢 (例: 86 歳) ではなく 1~19 の値で提供され、年齢情報が欠損したレセプトはないことが理解できます。なお、平均値などを算出する意味の薄いカテゴリカルな変数についても、平均値や標準偏差を (強引に) 計算している点にご留意ください。たとえば、表 2 の項番 12「男女区分」は最小値 1、最大値 2 で、性別を示す二値ですが、平均値 1.57 を表示しています (NDB の男女比がわかりますので、無意味というわけではありません)。

「各変数の項目値上位 50 サンプル」(「表番号 1-6」以降) については、より具体的に、各変数にどのような値が入っているのかを示しています。たとえば、表番号 1-7 は都道府県情報が入っています。NDB の生のデータでは、カラム名 TDFK の場所に 1~47 の数値が入っていることだけがわかりますが、「各変数の項目値上位 50 サンプル」を参照することにより、発行されたレセプト枚数をもっとも多い都道府県は東京であり、2 番目が北海道、もっとも少ない都道府県は鳥取であることがわかります。

このように、本コードブックは NDB を取り扱う際の辞書的な役割を果たします。

内容の詳細や著作権の取り扱いにつきましては、コードブック内の説明をご覧ください。

本コードブックが、先生方の NDB 研究に資することを願っております。

2019 年 7 月

奈良県立医科大学 公衆衛生学講座  
講師 野田 龍也