

「国立感染症研究所が提示する超過死亡数を 解釈する場合の注意点」についての編集部注記

栗原千絵子 (本号責任編集)

Editor's note to Suzumura's article "Points to note when interpreting on Excess Deaths presented by the National Institute of Infectious Diseases in Japan"

Chieko Kurihara (Responsible editor of Clinical Evaluation Vol.52 No.1)

標記の投稿論文*¹ (以下「本論文」) については、著者・査読者間で詳細な議論が交わされた。本論文での議論は、国立感染症研究所 (以下「感染研」) がレンジの形式で示される累積死亡数を算出している方法が世界標準と異なることを問題とし、累積超過死亡数—累積過少死亡数を新たに累積超過死亡数として算出すれば、世界標準に近い値になるとした。査読者はこの量が、統計的観点からどのような量になるかが明確でないとして疑念を示した。またそもそもレンジとしての基礎超過死亡数を足しこむという操作自体が統計的に不明確な統計量を生み出し問題であるとした。さらに統計的観点からレンジで示す超過死亡数の上限値が、観測された超過死亡数ということによいのか、という査読者の問いかけから何度も著者と査読者の間でやり取りがなされた。最終的に著者は5節に示されるような記述を選択したが、査読者はなお納得していない。このように受理に至るまでに、著者が示す考え方に対する疑義、残された課題が複

数あるが、「超過死亡数」の算出方法をめぐる議論は、感染研の提示する数値は社会的な影響力が大きいため、感染研において社会に対する説明責任があるとともに、今後専門家の中で議論を深める必要があると編集部において判断したことから、この点及び他の論点も含めオープン・クエスチョンとして読者諸賢の判断と議論に委ねるものとして掲載することとした。

また、本論文の脚注2 (編集部注, 本号52頁) に示すWHOの指標についての本論文著者の解釈に対し疑問が残った。この点について査読者より貴重な解説を得たので、以下に引用する。本来は記名で掲載すべき内容であるが、ブラインドで開始された査読の経緯において査読者の厚意により提供いただいた解説なので、匿名のまま引用する。この点は今後、著者又は関心のある読者がWHOに照会するなどして解決されることが望まれ、「超過死亡数」の算出方法とあわせて、関連情報、意見が寄せられることを願う。

*¹ 鈴村 泰. 国立感染症研究所が提示する超過死亡数を解釈する場合の注意点. 臨床評価. 2024 ; 52 (1) : 43-58.

査読者による解説： 信頼区間と予測区間について

本論文中に脚注2（編集部注，本号52頁）を付したWHOの指標は，その指標の意味を十分理解していないとデータの解釈を誤る可能性があるため，基礎的な概念を解説する．引用されたWHOの文献^{*2}中では「信頼区間」と記載されているが，「信頼区間」ではなく「予測区間」の意味で使われている可能性がある．

予測区間というのは，予測値ではなく観測値がどれほどの幅で動きうるかを示す指標である．このため，（超過死亡がないという仮説の下での）期待死亡数Eの予測値（予測死亡数）Yとその周りの予測区間（L, U）が与えられたときに，もしその区間上限Uと区間下限Lの中に観測死亡数Xが入っていれば，その観測値は超過死亡（ないし過少死亡）を示唆しないと解釈できる（この場合超過死亡数=0と考えてよい）．

したがって $X > U$ なら超過死亡と判定し，そのときの超過死亡数は $X - U$ となる．また得られたデータから超過死亡数として，もっともらしい値は観測超過死亡数 $X - Y$ である．ただしこの値はデータのばらつきを考慮していない．

このことは過少死亡の場合にも適用でき，その場合は $X - L$ が（マイナスの値としての）過少死亡数（+に置き換えるなら $L - X$ ），また $X - Y$ が（マイナス値としての）もっともらしい過少死亡数となる（統計的立場からは上限ではない）．

上の $X - U$ や $X - L$ は超過（過少）死亡数に関して言い過ぎのない限界の値である．

それと同様の意味で， $U \geq X \geq L$ の場合は，超過死亡数は0とセットされる．

しかし $X - U$ や $X - L$ は，超過死亡や過少死亡に対して，言い過ぎはないものの，控えめすぎる

という考え方もできる．したがって $X - E$ （ただしEはYの期待値）の予測区間として， $X - L \geq X - Y \geq X - U$ という考え方もできる．これが観測値Xと期待値Eの乖離度（超過死亡ないし過少死亡）を表す指標となる．つまり $X - U > 0$ なら超過死亡， $X - L < 0$ なら過少死亡と評価すればよい．この意味で，超過死亡の場合，超過死亡数は確率的な意味で最大 $X - L$ の可能性がある，過少死亡の場合，過少死亡数は確率的な意味で最大 $U - X$ の可能性がある．ただしここでは過少死亡数を+で表現したときの話である．

また，超過死亡しか考えないモデルでは，片側推測で考えることもできるが，その上で過少死亡も考えるというのは考え方の上で矛盾している．最初に両側推測か，片側推測かを明示する必要がある．

以上のような考え方を背景に，超過死亡の場合上限値を単に $X - Y$ で押さえてよいのか，と著者に質問した．超過死亡の場合 $X - Y$ は単に超過死亡数の観測値に過ぎず， $X - U > 0$ で初めて観測値 $X - U$ が統計的観点から超過死亡数を示していると判断できるが，その上限値は $X - L$ で押さえられるのである．CDCが上限を $X - Y$ で押さえていることは確認したが，これはリスク管理の観点からは不思議な押さえ方でもある．

例えば津波予測でいえば，Xが津波の高さの予測値でUが予測上限，Lが予測下限の場合に，XやLが堤防の高さHを超えたときに問題となる． $L > H$ は最低でも堤防を超えるという意味で問題になり， $X > H$ も予想値が堤防を超えるので問題となる．しかしXを想定する上限値と考えてよいのか？可能性としては最大高さUの津波が起こる可能性を考えておかなければならないはずである．Xを想定し $X < H$ だから大丈夫と判断して，実際の津波が $U > H (> X)$ だったときに「想定外」だったとは言えないはずである．

^{*2} World Health Organization. Global excess deaths associated with COVID-19 (modelled estimates) [cited 2024 Jun 1]. Available from : <https://www.who.int/data/sets/global-excess-deaths-associated-with-covid-19-modelled-estimates>

以上の考え方を背景にしてWHOの文献をみると、用語表で、観測値 Y と期待値 E 、超過死亡数 $\delta = Y - E$ が定義されており、この δ を推定することなので、WHOのいう信頼区間は予測区間と同じ意味合いではないかと思われる。

以上を踏まえ、本論文中に脚注2（編集部注、本号52頁）を付した部分のWHOの指標は、「信頼区間」ではなく「予測区間」と解釈し、「WHOドキュメントには信頼区間とあるが、他との整合性を考慮して信頼区間を予測区間と読み替えた」旨の解釈を示すことが適切ではないか考える。

一方、信頼区間であるとしたら、基礎モデルにポアソン分布を仮定し、さらにもしパンデミックが生じていれば、ポアソン分布のパラメータが大きくなるという overdispersion パラメータ θ を導入する。その点、他のアプローチと異なることになる。そして $\theta = 1$ なら超過も過少もなし、 $\theta > 1$ なら超過、 $\theta < 1$ なら過少と判定しようという立場ということになる。そうすると確かにパラメータ推測の話になるので、信頼区間を作ることになるかもしれないが、実は ベイズ推測 を使っているようで、そうであれば、「信頼区間」というより「信用区間」ということになる。しかしWHO

は信頼区間と称しており、実際の意図については確認を要する。

このようにWHOの方法については未解明なままである。信頼区間という言葉は使われているものの、その信頼上限 and / or 下限を超過死亡数の判定に使うという説明はなく、予測値の信頼区間からはずれたときに超過死亡あるいは過少死亡と判定するということは通常考えられない。信頼区間は期待値に関するものであり、予測値のばらつきに加えて観測値のばらつきを考慮したものが予測区間になるからである。そのうえでWHOの文書には予測区間に関する記述もない。ただ観測値を X 、期待値を E としたとき、 $X - E$ を超過／過少死亡数とするという定義が最初になされているだけである。

WHOは基礎分布としてポアソン分布を仮定し、超過分散パラメータ θ を導入した回帰モデルを用いているようなので、このモデルで $\theta > 1$ とみなされるとき超過分散、 $\theta < 1$ とみなされるとき過少分散と判定するというアプローチの可能性はあるが、Methodsの記述からそこまで読み取することはできない。

* * *