

図 19 データ区分別の補正差の標準偏差

続いて、図 20 は、各手法による補正結果が補正前と比較してどの程度異なるのか(単純指数との乖離度)を示したグラフである。また、手法間の補正結果の関係をみるため、補正された指数を 2 つの手法ごとに散布図の形で比較したものが、図 22 である。

ここで、単純指数との乖離度、また手法間の補正結果の乖離度を示すために用いる指標は、平均二乗誤差とよばれるものであり、次式のように示される。

$$\text{平均二乗誤差} = \frac{1}{m} \sum_{i=1}^m (\hat{\theta}_i - \theta_i)^2 \quad (5.1)$$

ここで、 m は地域の数を示し、 $\hat{\theta}_i$ は補正後の指数、 θ_i は補正前の単純指数である。元のデータから乖離しすぎることなく補正を行うという考えのもとでは、この平均二乗誤差の値はより小さい方が良いということになる。

① 単純指数との比較

データ区別に平均二乗誤差の値を比較すると、0-39 歳高血圧のすべての手法において最も高い値となっている。ついで、40-49 歳高血圧、50-59 歳高血圧、高血圧のデータ区分で高い値を示す結果となった。男性高血圧、女性高血圧、50-59 歳高血圧(LWA を除く)、60-69 歳高血圧、70 歳以上高血圧においては、ほぼゼロに近い値となっており、全体としてはあまり補正がなされていないといえる。つまり、全体として補正が大きくなされたデータ区分は、0-39 歳高血圧、40-49 歳高血圧、50-59 歳高血圧であった。

さらに、手法別に平均二乗誤差の値を比較すると、ほぼすべてのデータ区分において高い値を示した LWA の平均二乗誤差は、40-49 歳高血圧と 50-59 歳高血圧において著しく高い値を示している。また、0-39 歳高血圧、40-49 歳高血圧では、ベイジアンによる 3 つ手法の平均二乗誤差も高い値を示しているのに対し、50-59 歳高血圧ではほぼゼロに近く、LWA のみ高い値を示している。また、LWA による平均二乗誤差が高血圧を除くすべてのデータ区分において、大きいことが分かる。唯一、LWA の平均二乗誤差が他の手法に比べて低い値を示した高血圧では、LWA

が最も小さい値を示しており、他のベイジアン手法よりも低い値となっていた。このような傾向は他のデータ区分にみられないため、特徴的な結果だといえる。

また、LWA に次いで平均二乗誤差の値が大きい傾向にあったのは、経験ベイズ法による補正結果である。ただし、男性高血圧、女性高血圧、50-59 歳高血圧、60-69 歳高血圧、70 歳以上高血圧の 5 つのデータ区分においては、経験ベイズ法の平均二乗誤差の値は大きいとはいえず、階層ベイズ法との差もあまりないものといえる。

階層ベイズ法による 2 つのモデル間の比較では、高血圧を除きポアソンガンマモデルよりもポアソン対数正規モデルの方が平均二乗誤差の値が大きい傾向にあることがわかった。

② 手法間の補正結果の比較

上記で記述した平均二乗誤差は、全体としての単純指数との乖離度を示したグラフであった。ここでは、手法間の乖離度を示す平均二乗誤差を算出し、グラフ化した図 21 と、2 つの手法ごとに比較した散布図(図 22)より、手法間の補正結果を比較する。

図 21 より、最も精度が高いと考えられる階層ベイズと LWA との差は、40-49 歳高血圧、0-39 歳高血圧、50-59 歳高血圧、高血圧の 4 つのデータ区分において大きな値を示している。これらのデータ区分においては、補正結果に大きな差が生じていることがわかる。

また、反対に、男性高血圧、女性高血圧、60-69 歳高血圧、70 歳以上高血圧では平均二乗誤差が約 0.01 と非常に小さな値となっており、最も精度が良い手法と考えられる階層ベイズ法と LWA の補正結果に大きな差が生じていない。

さらに、3 つのベイジアン手法を比較してみると、ベイジアン手法の中では比較的簡便な方法である経験ベイズ法と最も精度が良いと考えられる階層ベイズの手法は高血圧、男性高血圧、女性高血圧、50-59 歳高血圧、60-69 歳高血圧、70 歳以上高血圧の 6 つのデータ区分において、あまり大きな差は生じていないことがわかる。反対に、0-39 歳高血圧、40-49 歳高血圧の 2 つのデータ区分において、両手法による平均二乗誤差の値は大きく、またその差も大きいことがわかる。これら 2 つのデータ区分においては、ベイジアン手法においても補正結果に差異が生じているといえる。なお、経験ベイズ法と 2 つの階層ベイズ法の比較では、特に経験ベイズ法とポアソン対数正規モデルを仮定した階層ベイズ法による平均二乗誤差に差が生じており、補正結果に差が生じるものといえる。

図 22 より、手法間の補正結果の関係をみるができる。手法の精度が高いと考えられる階層ベイズ法と LWA の間では、補正結果に最もばらつきが生じていることがわかる。また、経験ベイズ法と階層ベイズ法の間における、補正結果の方が LWA と階層ベイズ法よりもばらつきが少ないといえる。さらに、階層ベイズ法による 2 つのモデル間(ポアソンガンマモデルとポアソン対数正規モデル)では、補正結果に大きな差異はみられず、ほぼ一致しているようにみえる。

そこで、これら 3 つのベイジアンによる手法の結果をさらに詳細に比較するため、補正結果の差を期待件数との関係でみたものが、図 23 である。グラフより、期待件数が約 140~1,000 程度の範囲において両手法による補正差が 0.2 以上となっており、補正結果の間に差が生じることがわかる。

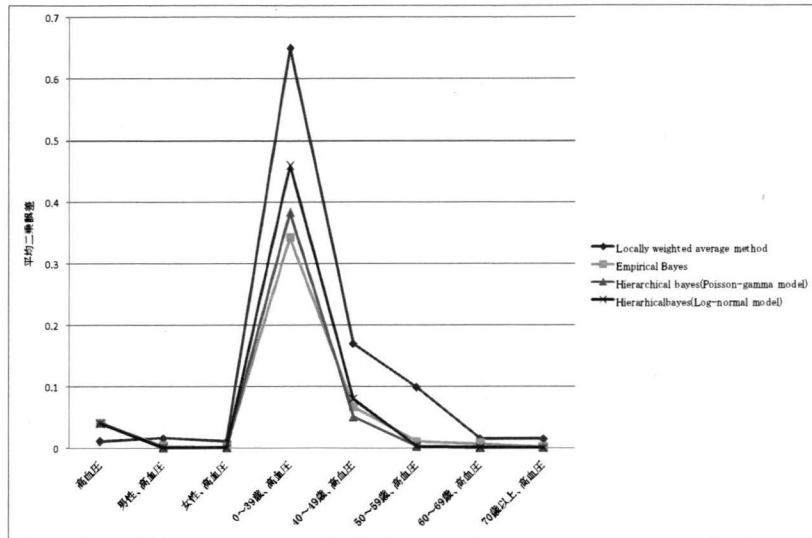


図 20 手法別の単純指数との平均二乗誤差

表 3 手法別の単純指数との平均二乗誤差

	Locally weighted average method	Empirical Bayes	Hierarchical bayes(Poisson-gamma model)	Hierarchical bayes(Log-normal model)
高血圧	0.012	0.041	0.041	0.040
男性、高血圧	0.017	0.004	0.000	0.001
女性、高血圧	0.012	0.002	0.000	0.000
0~39歳、高血圧	0.651	0.343	0.383	0.460
40~49歳、高血圧	0.171	0.067	0.051	0.080
50~59歳、高血圧	0.099	0.011	0.003	0.004
60~69歳、高血圧	0.016	0.008	0.001	0.001
70歳以上、高血圧	0.016	0.002	0.000	0.000

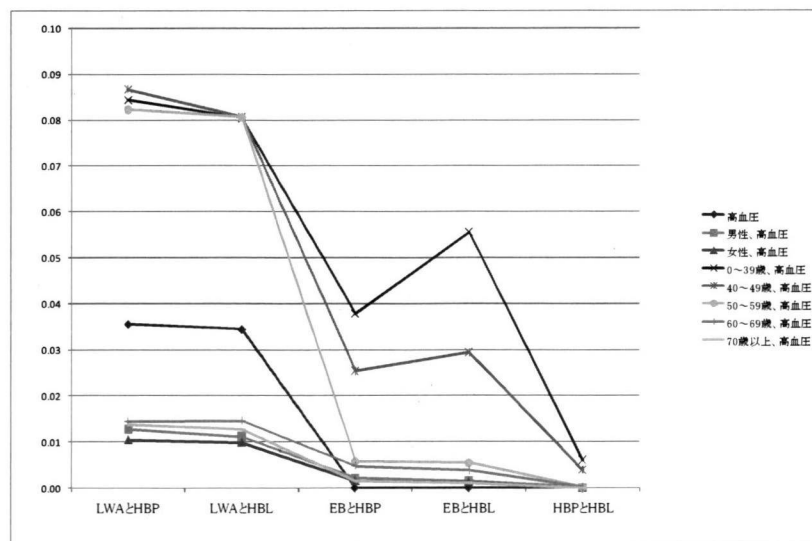


図 21 手法間の平均二乗誤差

表 4 手法間の平均二乗誤差

	LWAとHBP	LWAとHBL	EBとHBP	EBとHBL	HBPとHBL
高血圧	0.036	0.034	0.000	0.000	0.000
男性、高血圧	0.013	0.011	0.002	0.001	0.000
女性、高血圧	0.010	0.010	0.001	0.001	0.000
0～39歳、高血圧	0.084	0.080	0.038	0.056	0.006
40～49歳、高血圧	0.087	0.081	0.025	0.029	0.004
50～59歳、高血圧	0.082	0.081	0.006	0.005	0.000
60～69歳、高血圧	0.014	0.015	0.005	0.004	0.000
70歳以上、高血圧	0.014	0.013	0.001	0.001	0.000

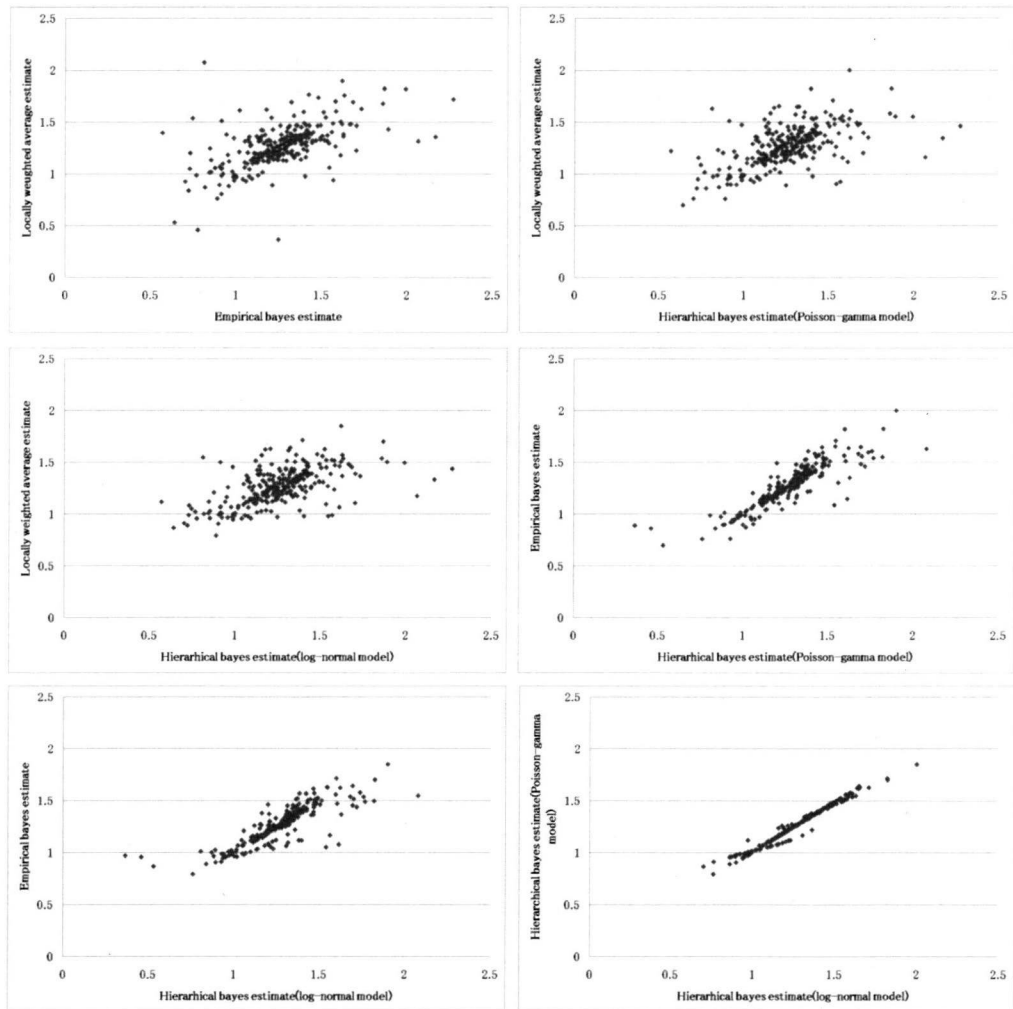


図 22 手法間の補正結果の比較

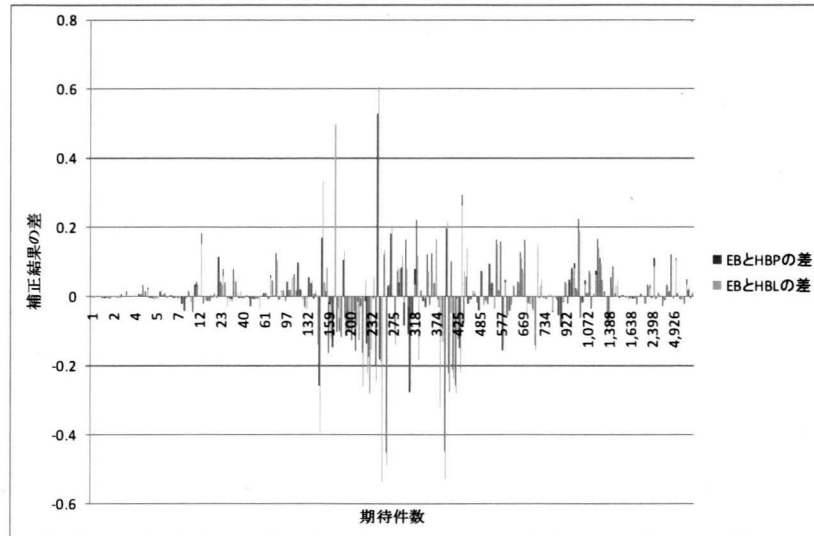


図 23 階層ベイズ法と経験ベイズ法の補正差と期待件数

また、階層ベイズ法による 2 つのモデルの比較では、DIC(Deviance Information Criterion)²³による比較が可能である。V. Gomez and Rubio et al.(2009)によれば、ベイジアンモデルを用いて推定を行うとき、最も良いモデルを選択するために DIC 値を計算する。DIC は同様のデータを用いて推定したときのみ用いることができる。また、MCMC 法による事後分布のサンプリングから得られる値であるため、今回用いた手法の中では、階層ベイズ法による 2 つの手法(ポアソンガンマモデルとポアソン対数正規モデル)のみ比較が可能となる。つまり、DIC は事後分布のサンプリングを用いて計算されるため、モーメント法を用いた経験ベイズ法の場合には算出が難しい。なお、経験ベイズ推定値の不確実性の評価については、別途 Butar and Lahiri(2003)によって提案されているが、本研究では複数の経験ベイズ推定値を算出している訳ではないため、それを行うには至らなかった。また、モデルの比較において、DIC の値は最も低い値を示すモデルが良いものとされることから、相対的なモデルの比較を可能とするものであることに注意が必要である。

表 5 より、データ区分別に DIC の値を確認すると、高血圧、60-69 歳高血圧、70 歳以上高血圧の 3 つのデータ区分においては、ポアソン対数正規モデルの方が DIC の値が小さく、モデルの当てはまりが良いといえる。また、男性高血圧、女性高血圧、0-39 歳高血圧、40-49 歳高血圧、50-59 歳高血圧の 5 つのデータ区分においては、ポアソンガンマモデルの方が DIC の値が小さく、モデルの当てはまりが良いという結果になった。つまり、期待件数が少ない傾向を示すデータ区分においては、ポアソンガンマモデルの方が良い結果にあり、反対に、比較的期待件数の多いデータ区分(高血圧、60-69 歳高血圧、70 歳以上高血圧)においては、ポアソン対数正規モデルの方は当てはまりが良いということがわかった。ただし、DIC の差は最も大きいものであっても、0-39 歳高血圧データでの約 4 程度となっており、著しい差ではない。

²³ 詳しくは、Lawson et al. (2003)あるいは、Spiegelhalter et al. (2002)を参照されたい。

表 5 階層ベイズ法によるモデルの DIC 値

	Poisson-gamma model	Log-normal model
高血圧	388.55	388.52
男性、高血圧	355.18	355.74
女性、高血圧	369.01	368.68
0～39歳、高血圧	144.01	147.75
40～49歳、高血圧	218.08	220.85
50～59歳、高血圧	296.34	297.41
60～69歳、高血圧	334.90	334.30
70歳以上、高血圧	372.28	372.04

最後に、上記の分析結果をデータ区分ごとにまとめると、以下のような結果となる。

① 高血圧

いずれの手法においても、単純指数との絶対値として示した補正差の最大値はおよそ 0.5 以下となっており、あまり大きな補正は行われなかった。

全体として単純指数との乖離度を示し、低い値であるほど良いとされる平均二乗誤差の値は、LWA が 0.012 と最も小さく、また、経験ベイズ法と階層ベイズ法は共におよそ 0.04 という値を示した。

手法間の差の程度を示す平均二乗誤差の値では、LWA と階層ベイズ法との間の値が 0.036 と比較的大きな値を示した。また、ベイジアン手法間における比較では、3つの手法間(経験ベイズ法と階層ベイズ法(ポアソンガンマモデルとポアソン対数正規モデル))で補正結果にほとんど差がなかった。さらに、階層ベイズ法による2つのモデルの比較では、DICの値からポアソン対数正規モデルの方が良いという結果が示された。ただし、その差は 0.03 と非常に小さな値であった。

② 男性高血圧

絶対値として示した補正差の最大値は、すべての手法で 0.4 以下となっており、大きな補正は行われていないことがわかった。単純指数との平均二乗誤差では、LWA が最も大きく 0.017、経験ベイズ法で 0.004、階層ベイズ法では非常に小さな値となった。

また、手法間の平均二乗誤差の値は、LWA と階層ベイズ法間においてポアソンガンマモデル 0.013、ポアソン対数正規モデル 0.011 という値となり、最も単純な方法である LWA と最も精度が高いとされる階層ベイズ法との間に大きな差は生じていないことがわかった。また、経験ベイズ法と階層ベイズ法間においても、大きな差は生じていないことがわかった。DICの値による階層ベイズ法の2つのモデルの比較では、ポアソンガンマモデルの方が良いという結果が得られたが、その差は非常にわずかな差であった。

③ 女性高血圧

絶対値として示した補正差の最大値は全体として非常に小さい傾向にあり、最も高い値で LWA の 0.29 であった。全体として、あまり補正が行われていないことがわかった。単純指数との平均二乗誤差では LWA の値が最も大きく、その他の手法においては非常に小さな値となった。

また、手法間の平均二乗誤差では、LWA と階層ベイズ法の間で 0.01 と非常に小さく、また、経験ベイズ法と階層ベイズ法、ポアソンガンマモデルとポアソン対数正規モデル間に差はほぼないことがわかった。DIC による階層ベイズ法の 2 つのモデルの比較では、ポアソン対数正規モデルの方が良いという結果になった。

④ 0-39 歳高血圧

絶対値として示した補正差の最大値は LWA で 2.65、経験ベイズ法で 1.82、階層ベイズ法(ポアソンガンマモデル)で 1.99、階層ベイズ法(ポアソン対数正規モデル)で 2.10 と、すべての手法において非常に大きな補正がなされていることがわかった。

また、単純指数との平均二乗誤差では、LWA で最も大きく 0.651、経験ベイズ法で 0.343、階層ベイズ法のポアソンガンマモデルで 0.383、階層ベイズ法のポアソン対数正規モデルで 0.46 と、経験ベイズ法が最も小さく、ついで階層ベイズ法のポアソンガンマモデルが小さい値となった。なお、経験ベイズ法による値とポアソンガンマモデルによる値にはわずかな差しか存在しなかった。

手法間の平均二乗誤差では、LWA と階層ベイズ法による 2 つのモデルが共におよそ 0.08 となり、最も大きな値を示した。両手法の補正結果は他の手法間の値よりも大きい結果となった。また、経験ベイズ法と階層ベイズ法による 2 つのモデル間では、ポアソンガンマモデルとの平均二乗誤差が 0.038、ポアソン対数正規モデルが 0.056 となり、階層ベイズ法の中でもポアソン対数正規モデルの方が経験ベイズ法との差が大きい傾向にあることがわかった。

また、DIC による比較では、ポアソンガンマモデルの方が良いという結果になった。

⑤ 40-49 歳高血圧

絶対値として示した補正差の最大値は、LWA で 1.27、経験ベイズ法で 0.52、階層ベイズ法のポアソンガンマモデルで 0.48、階層ベイズ法のポアソン対数正規モデルで 0.61 となっており、0-39 歳高血圧のデータ区分について、大きな値を示す結果となった。

単純指数との平均二乗誤差では、LWA の値が 0.171 と最も大きく、ついで経験ベイズ法による 0.067、階層ベイズ法のポアソンガンマモデルでは 0.051、階層ベイズ法のポアソン対数正規モデルでは 0.08 となり、階層ベイズ法のポアソン対数正規モデルが最も小さな値となった。

さらに、手法間の平均二乗誤差では、LWA と階層ベイズ法の 2 つのモデル間でおよそ 0.08 と最も大きな値となった。また、経験ベイズ法と階層ベイズ法の比較ではポアソン対数正規モデルとの平均二乗誤差の方がポアソンガンマモデルよりも大きい値を示していた。DIC による比較では、階層ベイズ法による手法のうち、ポアソンガンマモデルの方が良いという結果になった。

⑥ 50-59 歳高血圧

絶対値とした補正差の最大値は、LWA で 1.04、経験ベイズ法で 0.29、階層ベイズ法のポアソンガンマモデルで 0.14、階層ベイズ法のポアソン対数正規モデルで 0.16 と、0-39 歳高血圧、40-49 歳高血圧のデータ区分について比較的大きな補正がなされていることがわかった。

また、単純指数との平均二乗誤差では LWA で 0.09、経験ベイズ法で 0.01、階層ベイズ法のポアソンガンマモデルで 0.003、階層ベイズ法のポアソン対数正規モデルで 0.004 となっ

ており、LWA が最も大きく、階層ベイズ法による補正が最も良いということが示唆された。

さらに、手法間の平均二乗誤差では LWA と階層ベイズ法による 2 つのモデル間の値は大きくなっており、およそ 0.08 であった。一方、階層ベイズ法と経験ベイズ法間の値は 0.006 前後と非常に小さな値となった。

DIC による階層ベイズ法の 2 つのモデルの比較では、ポアソンガンマモデルが良いという結果になった。

⑦ 60-69 歳高血圧

絶対値で示した補正差の最大値は、LWA で 0.31、経験ベイズ法で 0.31、階層ベイズ法のポアソンガンマモデルで 0.08、ポアソン対数正規モデルで 0.12 といずれも著しく大きな補正はなされていないことがわかった。

単純指数との平均二乗誤差では、最大で LWA が 0.016、最小で階層ベイズ法の 2 つのモデルが 0.001 とすべての手法において小さな値を示している。また、手法間の平均二乗誤差でもすべての手法間で 0.015 以下となっており、手法間で差が生じていないことがわかった。

⑧ 70 歳以上高血圧

絶対値として示した補正差の最大値は、LWA で 0.49、経験ベイズ法で 0.16、階層ベイズ法のポアソンガンマモデルで 0.04、階層ベイズ法のポアソン対数正規モデルで 0.07 となり、すべての手法において大きな補正はなされていないことがわかった。

また、単純指数との平均二乗誤差では、LWA で最大の 0.016、3 つのベイジアン手法では共に 0.002 以下を示した。さらに、手法間の平均二乗誤差では、LWA と階層ベイズ法の値が最大となっており、0.014 前後であった。したがって、手法によって補正結果に大きな差は生じていないということがわかった。

最後に、表 6 から表 13 の指数一覧表を確認すると、0-39 歳高血圧のデータ区分において実績件数がゼロとなる市町村がいくつか存在していることがわかる。これらの指数は単純指数として示すと、当然ながらゼロと算出される。しかし、各種の補正手法により算出された値はゼロではなく、実数値が算出されている。したがって、McEwin and Elazar(2006)も指摘しているように、標本が得られないような場合にも SAE の手法は役立つものといえる。

また、経験ベイズ法の場合には、期待件数が少ないときに補正式中の ϕ 二乗が負値となり、ゼロで打ち切りとなる市町村がいくつかみられた。この場合には補正はなされず、結果として補正前の単純指数の値が返される。これは、モーメント法を用いた経験ベイズ法の欠点だと考えられる。

表 6 山形県 高血圧データの補正結果

	市町村	データ				Locally weighted average	Hierarchical bayes		
		被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model
1	山形市	22,813,756	1,230	854	1.440	1.194	1.219	1.219	1.218
2	米沢市	8,276,604	2,223	1,510	1.472	1.424	1.408	1.407	1.406
3	鶴岡市	15,145,025	16,594	13,618	1.219	1.241	1.156	1.157	1.157
4	酒田市	12,351,192	7,013	4,981	1.408	1.405	0.978	0.979	0.980
5	新庄市	4,607,092	2,861	2,398	1.193	1.275	1.134	1.134	1.135
6	寒河江市	4,094,790	963	692	1.391	1.354	1.194	1.194	1.194
7	上山市	3,645,684	1,506	1,204	1.251	1.277	1.192	1.192	1.193
8	村山市	2,919,189	835	501	1.668	1.340	1.168	1.167	1.169
9	長井市	2,869,720	1,596	1,165	1.370	1.412	1.275	1.270	1.270
10	天童市	6,052,830	760	638	1.192	1.395	1.241	1.241	1.240
11	東根市	4,380,985	802	637	1.258	1.265	1.333	1.334	1.334
12	尾花沢市	2,604,504	711	567	1.254	1.221	1.245	1.245	1.245
13	南陽市	3,500,126	916	674	1.358	1.377	1.340	1.337	1.336
14	山辺町	1,389,794	480	426	1.127	1.148	1.427	1.431	1.427
15	中山町	1,149,701	900	752	1.196	1.170	1.379	1.385	1.380
16	河北町	1,968,918	513	433	1.186	1.292	1.251	1.251	1.251
17	西川町	708,624	821	681	1.206	1.178	1.623	1.640	1.627
18	朝日町	1,052,576	370	270	1.371	1.116	1.196	1.196	1.197
19	大江町	1,020,201	439	355	1.236	1.252	1.257	1.259	1.258
20	大石田町	1,002,071	523	382	1.369	1.214	1.253	1.255	1.254
21	金山町	840,455	2,668	2,240	1.191	1.246	1.137	1.137	1.143
22	最上町	1,406,370	2,075	1,781	1.165	1.229	1.195	1.199	1.200
23	舟形町	740,999	2,170	1,708	1.271	1.329	1.186	1.192	1.193
24	真室川町	1,235,430	4,120	3,321	1.240	1.164	1.204	1.210	1.210
25	大蔵村	526,288	3,290	2,464	1.335	1.259	1.326	1.358	1.350
26	蛙川村	675,472	1,847	1,484	1.245	1.292	1.226	1.239	1.239
27	戸沢村	709,919	2,633	1,966	1.339	1.318	1.335	1.360	1.353
28	高島町	2,723,385	1,309	983	1.332	1.429	1.468	1.467	1.464
29	川西町	1,963,220	780	510	1.531	1.385	1.370	1.367	1.364
30	小国町	938,616	574	507	1.132	1.078	1.360	1.354	1.350
31	白鷹町	1,662,003	1,532	1,638	0.936	1.072	1.335	1.330	1.328
32	飯豊町	818,958	1,111	1,093	1.017	0.915	1.515	1.513	1.503
33	三川町	864,801	10,352	8,949	1.157	1.120	1.127	1.141	1.146
34	庄内町	2,804,193	7,280	7,446	0.978	0.990	0.939	0.943	0.948
35	遊佐町	1,816,885	2,754	2,433	1.132	1.155	1.018	1.025	1.030

データ出典:「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 7 山形県 男性 高血圧データ補正結果

データ					Locally weighted average	Empirical bayes	Hierarchical bayes		
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model	
1	山形市	40,168	6,702	5,462	1.227	1.250	1.227	1.227	1.227
2	米沢市	14,858	2,728	1,984	1.375	1.391	1.367	1.372	1.371
3	鶴岡市	27,314	4,130	3,543	1.166	1.133	1.162	1.166	1.167
4	酒田市	21,856	2,848	2,938	0.970	0.984	0.973	0.973	0.975
5	新庄市	8,360	1,102	982	1.122	1.160	1.127	1.126	1.128
6	寒河江市	7,552	1,193	1,010	1.181	1.237	1.190	1.184	1.184
7	上山市	6,680	1,115	922	1.209	1.228	1.212	1.210	1.210
8	村山市	5,403	819	734	1.115	1.220	1.141	1.121	1.124
9	長井市	5,328	899	703	1.278	1.305	1.319	1.277	1.274
10	天童市	11,122	1,653	1,395	1.185	1.131	1.191	1.186	1.187
11	東根市	8,135	1,406	1,049	1.340	1.201	1.321	1.337	1.334
12	尾花沢市	4,893	718	623	1.153	1.250	1.172	1.158	1.160
13	南陽市	6,379	1,015	800	1.269	1.315	1.313	1.267	1.265
14	山辺町	2,532	474	353	1.342	1.163	1.299	1.334	1.326
15	中山町	2,143	402	296	1.359	1.282	1.304	1.347	1.338
16	河北町	3,646	619	504	1.228	1.253	1.228	1.229	1.229
17	西川町	1,295	363	212	1.712	1.314	1.466	1.649	1.618
18	朝日町	1,984	314	269	1.167	1.366	1.194	1.175	1.178
19	大江町	1,882	314	263	1.193	1.243	1.209	1.198	1.200
20	大石田町	1,899	275	232	1.183	1.165	1.205	1.191	1.192
21	金山町	1,602	183	171	1.068	1.126	1.110	1.096	1.113
22	最上町	2,647	346	309	1.120	1.132	1.134	1.131	1.138
23	舟形町	1,401	208	180	1.155	1.245	1.161	1.168	1.173
24	真室川町	2,299	313	279	1.122	1.131	1.136	1.134	1.140
25	大蔵村	971	165	108	1.528	1.131	1.344	1.460	1.428
26	鮭川村	1,277	189	147	1.284	1.222	1.234	1.276	1.269
27	戸沢村	1,355	232	161	1.442	1.191	1.329	1.408	1.387
28	高島町	5,216	912	637	1.432	1.326	1.384	1.423	1.416
29	川西町	3,688	633	480	1.319	1.357	1.340	1.314	1.309
30	小国町	1,646	375	275	1.363	1.381	1.353	1.351	1.341
31	白鷹町	3,203	574	425	1.349	1.343	1.350	1.341	1.335
32	飯豊町	1,513	308	207	1.486	1.349	1.375	1.453	1.433
33	三川町	1,631	236	207	1.142	1.051	1.114	1.156	1.163
34	庄内町	5,168	625	670	0.933	1.041	0.951	0.947	0.959
35	遊佐町	3,267	409	424	0.965	0.915	0.985	0.985	1.000

データ出典:「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 8 山形県 女性 高血圧データ補正結果

データ					Locally weighted average	Empirical bayes	Hierarchical bayes		
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model	
1	山形市	47,916	9,892	8,156	1.213	1.234	1.214	1.213	1.213
2	米沢市	17,098	4,285	2,996	1.430	1.414	1.427	1.428	1.427
3	鶴岡市	31,161	6,222	5,406	1.151	1.111	1.149	1.152	1.152
4	酒田市	25,832	4,432	4,508	0.983	0.994	0.986	0.986	0.988
5	新庄市	9,428	1,652	1,450	1.139	1.152	1.170	1.143	1.145
6	寒河江市	8,258	1,668	1,388	1.201	1.303	1.208	1.203	1.204
7	上山市	7,396	1,553	1,317	1.179	1.259	1.189	1.181	1.183
8	村山市	5,868	1,256	1,046	1.201	1.236	1.210	1.203	1.204
9	長井市	5,752	1,271	1,004	1.266	1.346	1.302	1.266	1.266
10	天童市	12,248	2,467	1,926	1.281	1.187	1.278	1.280	1.280
11	東根市	8,780	1,884	1,415	1.332	1.302	1.321	1.330	1.329
12	尾花沢市	5,163	1,129	861	1.311	1.322	1.299	1.310	1.308
13	南陽市	7,135	1,618	1,166	1.387	1.319	1.391	1.383	1.382
14	山辺町	2,834	756	501	1.508	1.215	1.431	1.492	1.483
15	中山町	2,296	561	397	1.414	1.409	1.357	1.403	1.395
16	河北町	3,956	887	700	1.267	1.295	1.264	1.268	1.267
17	西川町	1,441	472	289	1.635	1.359	1.471	1.596	1.575
18	朝日町	2,080	446	369	1.210	1.416	1.226	1.216	1.218
19	大江町	2,057	488	374	1.305	1.280	1.286	1.301	1.299
20	大石田町	1,970	436	334	1.304	1.261	1.284	1.301	1.298
21	金山町	1,643	297	255	1.166	1.164	1.193	1.180	1.187
22	最上町	2,783	554	443	1.249	1.196	1.208	1.251	1.252
23	舟形町	1,460	305	252	1.208	1.325	1.198	1.217	1.219
24	真室川町	2,471	508	402	1.265	1.210	1.210	1.266	1.265
25	大蔵村	1,061	205	162	1.266	1.107	1.203	1.268	1.266
26	鮭川村	1,331	250	208	1.203	1.274	1.198	1.213	1.218
27	戸沢村	1,386	291	221	1.315	1.230	1.211	1.310	1.305
28	高島町	5,299	1,311	873	1.501	1.495	1.472	1.492	1.486
29	川西町	3,892	963	685	1.405	1.451	1.404	1.398	1.394
30	小国町	1,978	541	399	1.355	1.374	1.377	1.348	1.343
31	白鷹町	3,214	735	557	1.319	1.495	1.352	1.316	1.314
32	飯豊町	1,649	472	302	1.561	1.409	1.475	1.530	1.514
33	三川町	1,708	338	300	1.125	1.097	1.104	1.141	1.151
34	庄内町	5,659	907	967	0.938	1.094	0.955	0.950	0.960
35	遊佐町	3,748	702	669	1.049	0.914	1.052	1.060	1.070

データ出典:「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 9 山形県 0-39 歳 高血圧データ補正結果

		データ				Locally weighted average	Empirical bayes	Hierarchical bayes	
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model	
1	山形市	19,432	53	36	1.475	1.320	1.361	1.364	1.221
2	米沢市	6,548	19	12	1.564	1.339	1.564	1.306	1.170
3	鶴岡市	12,080	12	23	0.532	0.642	0.532	0.702	0.868
4	酒田市	9,420	17	18	0.971	1.002	0.971	0.991	1.011
5	新庄市	4,499	6	8	0.719	0.958	0.884	0.899	0.968
6	寒河江市	3,462	2	6	0.311	0.704	0.929	0.765	0.914
7	上山市	2,781	8	5	1.543	1.704	1.228	1.204	1.108
8	村山市	2,238	2	4	0.480	0.854	1.020	0.875	0.962
9	長井市	2,326	2	4	0.460	0.778	0.460	0.864	0.959
10	天童市	5,688	8	11	0.758	0.924	1.003	0.901	0.968
11	東根市	3,915	6	7	0.825	0.730	1.052	0.952	0.991
12	尾花沢市	2,156	3	4	0.749	0.906	1.072	0.954	0.996
13	南陽市	3,097	5	6	0.872	0.819	0.872	0.977	1.003
14	山辺町	1,004	3	2	1.602	1.612	1.183	1.118	1.068
15	中山町	933	2	2	1.150	1.380	1.142	1.049	1.038
16	河北町	1,466	3	3	1.100	0.855	1.135	1.045	1.032
17	西川町	377	0	1	0.000	0.923	1.100	0.973	1.009
18	朝日町	813	3	2	1.975	0.733	1.203	1.160	1.085
19	大江町	734	0	1	0.000	1.546	1.063	0.906	0.981
20	大石田町	845	4	2	2.521	0.848	1.248	1.233	1.122
21	金山町	832	1	2	0.636	0.845	1.015	0.988	1.009
22	最上町	1,317	5	2	2.031	0.944	1.382	1.220	1.120
23	舟形町	585	0	1	0.000	1.570	0.942	0.928	0.990
24	真室川町	1,126	4	2	1.903	1.534	1.320	1.184	1.097
25	大蔵村	503	3	1	3.218	0.570	1.398	1.224	1.119
26	鮭川村	616	2	1	1.733	1.019	1.215	1.107	1.062
27	戸沢村	635	0	1	0.000	1.060	0.931	0.925	0.988
28	高島町	2,442	6	5	1.308	1.079	1.308	1.110	1.066
29	川西町	1,468	1	3	0.366	1.249	0.366	0.893	0.972
30	小国町	446	1	1	1.232	1.183	1.232	1.049	1.039
31	白鷹町	1,325	4	2	1.616	1.021	1.616	1.150	1.080
32	飯豊町	605	1	1	0.893	1.213	0.893	1.017	1.028
33	三川町	691	2	1	1.541	0.747	1.541	1.090	1.054
34	庄内町	2,167	4	4	0.988	0.769	0.988	1.019	1.024
35	遊佐町	1,324	2	2	0.809	0.915	0.809	0.990	1.013

データ出典:「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 10 山形県 40-49 歳 高血圧データ補正結果

データ					Locally weighted average	Empirical bayes	Hierarchical bayes		
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model	
1	山形市	5,442	183	132	1.384	1.439	1.387	1.391	1.390
2	米沢市	2,093	83	51	1.620	1.859	1.681	1.585	1.542
3	鶴岡市	3,983	127	98	1.301	1.220	1.247	1.321	1.333
4	酒田市	2,976	83	73	1.142	1.101	1.154	1.193	1.232
5	新庄市	1,522	43	37	1.158	1.341	1.158	1.241	1.278
6	寒河江市	1,019	26	25	1.041	1.240	1.270	1.186	1.252
7	上山市	866	40	21	1.888	1.525	1.548	1.711	1.630
8	村山市	708	16	17	0.929	1.582	1.268	1.159	1.239
9	長井市	703	22	17	1.274	1.733	1.630	1.353	1.369
10	天童市	1,598	44	39	1.131	1.084	1.274	1.216	1.262
11	東根市	1,177	46	29	1.599	1.352	1.472	1.550	1.508
12	尾花沢市	799	41	20	2.092	1.395	1.601	1.823	1.717
13	南陽市	1,029	41	25	1.625	1.582	1.702	1.560	1.520
14	山辺町	344	11	8	1.315	1.326	1.381	1.396	1.401
15	中山町	250	7	6	1.141	1.629	1.364	1.351	1.370
16	河北町	488	14	12	1.170	1.354	1.347	1.317	1.347
17	西川町	145	7	4	1.957	1.890	1.433	1.557	1.506
18	朝日町	295	8	7	1.107	1.386	1.355	1.329	1.359
19	大江町	267	12	7	1.839	1.304	1.448	1.574	1.517
20	大石田町	292	10	7	1.400	1.402	1.394	1.431	1.422
21	金山町	303	9	7	1.208	1.069	1.208	1.361	1.374
22	最上町	471	17	12	1.469	1.259	1.469	1.458	1.440
23	舟形町	202	9	5	1.821	1.993	1.821	1.554	1.499
24	真室川町	356	14	9	1.606	1.585	1.606	1.514	1.473
25	大蔵村	216	11	5	2.078	0.813	2.078	1.632	1.551
26	鮭川村	246	7	6	1.166	1.208	1.166	1.361	1.382
27	戸沢村	229	10	6	1.768	1.425	1.768	1.544	1.492
28	高島町	759	45	19	2.417	1.621	1.900	2.002	1.853
29	川西町	529	20	13	1.520	1.684	1.696	1.485	1.457
30	小国町	164	6	4	1.494	2.273	1.720	1.463	1.440
31	白鷹町	439	25	11	2.310	1.868	1.825	1.825	1.704
32	飯豊町	198	10	5	2.049	1.634	1.760	1.610	1.540
33	三川町	225	9	6	1.610	1.145	1.201	1.495	1.465
34	庄内町	776	12	19	0.630	1.079	1.048	0.973	1.121
35	遊佐町	468	10	12	0.866	0.959	1.123	1.180	1.262

データ出典:「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 11 山形県 50-59 歳 高血圧データ補正結果

		データ				Locally weighted average	Empirical bayes	Hierarchical bayes	
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model	
1	山形市	10,585	1,167	921	1.267	1.383	1.288	1.268	1.268
2	米沢市	4,112	526	355	1.483	1.673	1.483	1.475	1.473
3	鶴岡市	8,424	806	726	1.110	1.066	1.101	1.116	1.117
4	酒田市	6,701	516	581	0.889	0.903	0.896	0.903	0.908
5	新庄市	2,599	246	225	1.093	1.295	1.129	1.112	1.115
6	寒河江市	2,092	229	181	1.267	1.381	1.314	1.273	1.271
7	上山市	1,804	226	155	1.455	1.299	1.365	1.440	1.435
8	村山市	1,577	175	135	1.299	1.467	1.325	1.303	1.300
9	長井市	1,483	179	127	1.406	1.560	1.406	1.395	1.391
10	天童市	3,283	367	284	1.291	1.052	1.317	1.293	1.293
11	東根市	2,308	298	200	1.492	1.421	1.383	1.477	1.472
12	尾花沢市	1,638	196	140	1.396	1.281	1.348	1.387	1.383
13	南陽市	1,913	264	165	1.597	1.512	1.597	1.568	1.562
14	山辺町	747	103	65	1.589	1.122	1.366	1.529	1.518
15	中山町	593	70	51	1.366	1.372	1.336	1.356	1.349
16	河北町	1,014	138	87	1.580	1.534	1.374	1.533	1.524
17	西川町	272	24	23	1.031	2.067	1.318	1.162	1.178
18	朝日町	561	68	48	1.413	1.141	1.341	1.389	1.380
19	大江町	525	56	45	1.248	1.283	1.325	1.271	1.269
20	大石田町	680	74	58	1.273	1.144	1.326	1.284	1.284
21	金山町	572	54	49	1.098	0.952	1.186	1.162	1.167
22	最上町	889	100	76	1.309	1.089	1.282	1.312	1.308
23	舟形町	476	48	41	1.179	1.328	1.224	1.224	1.227
24	真室川町	761	88	65	1.347	1.403	1.298	1.341	1.338
25	大蔵村	333	45	29	1.574	0.987	1.348	1.475	1.458
26	鮭川村	465	57	40	1.430	1.134	1.318	1.396	1.387
27	戸沢村	481	70	41	1.706	1.157	1.424	1.587	1.571
28	高畠町	1,664	242	143	1.696	1.325	1.696	1.651	1.645
29	川西町	1,244	152	106	1.428	1.344	1.428	1.411	1.408
30	小国町	366	43	32	1.358	2.168	1.358	1.346	1.336
31	白鷹町	937	121	80	1.506	1.612	1.506	1.471	1.464
32	飯豊町	432	64	37	1.739	1.481	1.739	1.598	1.581
33	三川町	544	54	47	1.157	0.877	1.060	1.206	1.211
34	庄内町	1,714	102	148	0.691	0.891	0.763	0.763	0.795
35	遊佐町	1,129	75	97	0.771	0.724	0.840	0.864	0.891

データ出典：「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 12 山形県 60-69 歳 高血圧データ補正結果

データ					Locally weighted average	Empirical bayes	Hierarchical bayes		
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model	
1	山形市	19,447	4,446	3,354	1.326	1.349	1.328	1.326	1.326
2	米沢市	7,262	1,836	1,253	1.466	1.428	1.471	1.462	1.460
3	鶴岡市	12,502	2,663	2,154	1.236	1.190	1.232	1.238	1.239
4	酒田市	10,784	1,829	1,849	0.989	0.990	0.994	0.996	1.001
5	新庄市	3,535	744	605	1.230	1.203	1.282	1.238	1.240
6	寒河江市	3,477	803	599	1.340	1.343	1.340	1.340	1.339
7	上山市	3,168	721	547	1.318	1.368	1.330	1.319	1.320
8	村山市	2,295	489	394	1.240	1.445	1.304	1.249	1.253
9	長井市	2,471	581	425	1.367	1.382	1.454	1.366	1.363
10	天童市	5,194	1,165	894	1.303	1.179	1.320	1.305	1.306
11	東根市	3,764	927	650	1.426	1.265	1.381	1.423	1.420
12	尾花沢市	1,835	414	318	1.303	1.517	1.329	1.308	1.310
13	南陽市	2,832	752	485	1.552	1.399	1.492	1.539	1.529
14	山辺町	1,179	297	202	1.469	1.333	1.369	1.451	1.440
15	中山町	974	249	168	1.486	1.505	1.368	1.463	1.449
16	河北町	1,718	407	296	1.377	1.424	1.351	1.375	1.370
17	西川町	600	177	105	1.693	1.629	1.385	1.610	1.573
18	朝日町	795	169	136	1.239	1.447	1.324	1.261	1.269
19	大江町	817	186	142	1.310	1.493	1.335	1.317	1.318
20	大石田町	670	161	114	1.407	1.389	1.349	1.394	1.384
21	金山町	573	127	98	1.298	1.327	1.303	1.312	1.313
22	最上町	968	233	167	1.394	1.381	1.313	1.387	1.380
23	舟形町	546	116	94	1.233	1.544	1.299	1.265	1.272
24	真室川町	913	205	158	1.301	1.313	1.303	1.310	1.310
25	大蔵村	358	86	61	1.399	1.227	1.307	1.382	1.370
26	鮭川村	443	117	76	1.538	1.535	1.315	1.481	1.454
27	戸沢村	500	125	86	1.448	1.410	1.311	1.421	1.405
28	高島町	2,101	561	361	1.555	1.567	1.489	1.537	1.526
29	川西町	1,508	392	259	1.514	1.665	1.480	1.495	1.484
30	小国町	913	219	159	1.378	1.477	1.467	1.374	1.367
31	白鷹町	1,304	312	222	1.407	1.705	1.467	1.399	1.393
32	飯豊町	648	178	112	1.591	1.621	1.482	1.536	1.508
33	三川町	645	143	111	1.284	1.247	1.213	1.298	1.302
34	庄内町	2,288	381	395	0.965	1.130	0.987	0.995	1.018
35	遊佐町	1,474	254	254	1.002	0.979	1.026	1.042	1.070

データ出典:「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

表 13 山形県 70 歳以上 高血圧データ補正結果

データ					Locally weighted average	Empirical bayes	Hierarchical bayes	
市町村	被保険者数	実績件数	期待件数	単純指数		Poisson-gamma model (モーメント法)	Poisson-gamma model	Log-normal model
1	山形市	33,178	10,745	9,175	1.171	1.184	1.171	1.171
2	米沢市	11,941	4,549	3,310	1.374	1.361	1.370	1.371
3	鶴岡市	21,486	6,744	5,948	1.134	1.101	1.131	1.134
4	酒田市	17,807	4,835	4,926	0.982	0.998	0.985	0.985
5	新庄市	5,633	1,715	1,558	1.101	1.113	1.112	1.104
6	寒河江市	5,760	1,801	1,587	1.135	1.240	1.141	1.137
7	上山市	5,457	1,673	1,511	1.107	1.190	1.116	1.110
8	村山市	4,453	1,393	1,230	1.133	1.130	1.140	1.135
9	長井市	4,097	1,386	1,134	1.223	1.279	1.245	1.222
10	天童市	7,607	2,536	2,094	1.211	1.175	1.210	1.212
11	東根市	5,751	2,013	1,579	1.275	1.237	1.268	1.274
12	尾花沢市	3,628	1,193	1,002	1.190	1.222	1.191	1.192
13	南陽市	4,643	1,571	1,286	1.222	1.259	1.242	1.222
14	山辺町	2,092	816	577	1.414	1.150	1.366	1.401
15	中山町	1,689	635	466	1.364	1.295	1.320	1.353
16	河北町	2,916	944	806	1.171	1.196	1.175	1.173
17	西川町	1,342	627	369	1.701	1.207	1.546	1.656
18	朝日町	1,600	512	444	1.152	1.409	1.164	1.156
19	大江町	1,596	548	443	1.238	1.188	1.227	1.238
20	大石田町	1,382	462	386	1.198	1.181	1.197	1.200
21	金山町	965	289	270	1.070	1.123	1.116	1.088
22	最上町	1,785	545	495	1.101	1.110	1.120	1.110
23	舟形町	1,052	340	292	1.165	1.192	1.139	1.172
24	真室川町	1,614	510	447	1.141	1.088	1.134	1.148
25	大蔵村	622	225	174	1.296	1.111	1.157	1.280
26	鮭川村	838	256	232	1.103	1.182	1.124	1.119
27	戸沢村	896	318	248	1.282	1.151	1.163	1.273
28	高島町	3,549	1,369	983	1.392	1.383	1.376	1.385
29	川西町	2,831	1,031	784	1.315	1.334	1.318	1.310
30	小国町	1,735	647	479	1.351	1.284	1.341	1.342
31	白鷹町	2,412	847	667	1.269	1.310	1.287	1.267
32	飯豊町	1,279	527	355	1.485	1.297	1.411	1.461
33	三川町	1,234	366	342	1.070	1.051	1.065	1.084
34	庄内町	3,882	1,033	1,072	0.964	1.077	0.979	0.972
35	遊佐町	2,620	770	728	1.058	0.917	1.057	1.065

データ出典：「国、都道府県の医療費適正化計画の重点対象の発見に関する研究」のデータより著者作成
 ※被保険者数とは、該当するデータ区分のみの被保険者数を指す。

5.1.2. 考察

以上の分析結果より、全体として補正が大きくなされたデータ区分は、0-39 歳高血圧、40-49 歳高血圧、50-59 歳高血圧であり、これらのデータ区分は他のデータ区分に比べ期待件数が少ない傾向にあることがわかった。これら 3 つの区分では、すべての手法で大きく補正がなされていることから、補正の必要性が高いデータ区分だと考えられる。例として、0-39 歳高血圧の補正結果を地図化したものが、図 24 である。精度の高い手法(右に位置する手法ほど精度が高い)による補正結果ほど指数の変化のグラデーションが滑らかになっていることが視覚的に確認することができる。手法により補正結果の大きく異なるデータ区分の場合には、マッピングした結果も大きく異なっていることから、政策判断材料として用いる際の補正の重要性が示唆される。

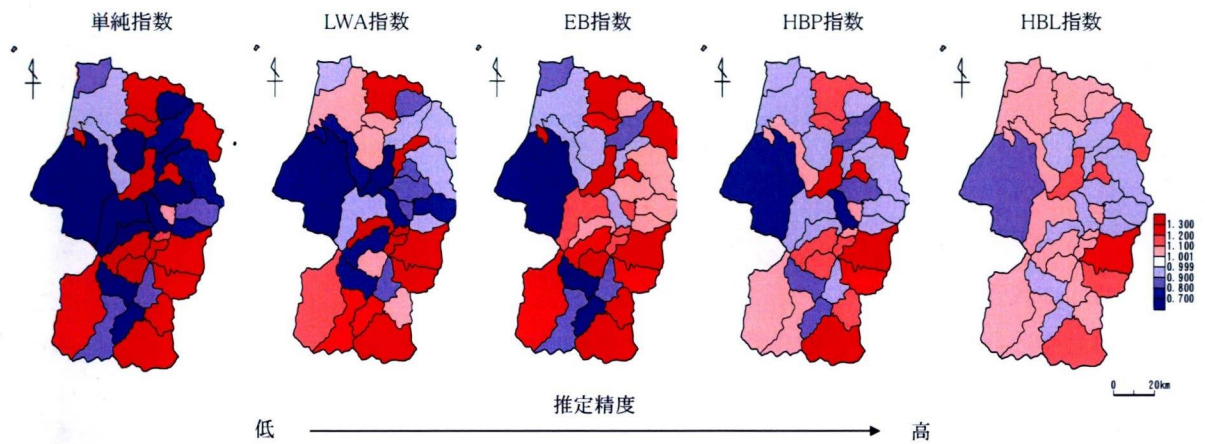


図 24 0-39 歳高血圧、受診指数の補正結果

一方、上記 3 つのデータ区分以外(高血圧、男性高血圧、女性高血圧、60-69 歳高血圧、70 歳以上高血圧の 5 つ)においては、各手法による補正結果の差に著しい差異は観察されず、上記 3 つのデータ区分に比べて、補正の必要性が低いものと考えられる。例として、高血圧と男性高血圧の補正結果を地図化したものが、図 25、図 26 である。上記に示した 0-39 歳高血圧の受診指数ほどは補正がなされていないことが視覚的に確認できる。また、高血圧のデータ区分では、経験ベイズ法と階層ベイズ法による補正結果は酷似していることが確認できる。さらに、男性高血圧のデータ区分では、殆ど補正がなされておらず、地図の様子もほとんど変化していないことが視覚的に確認できる。この点は、上記の 0-39 歳のデータ区分に関する補正結果と大きな異なっており、特徴的である。

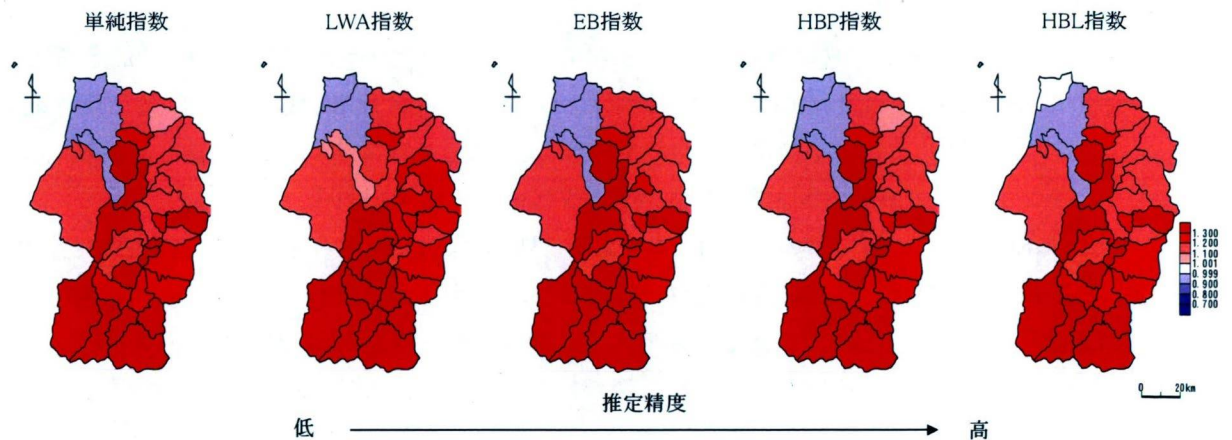


図 25 高血圧、受診指数の補正結果

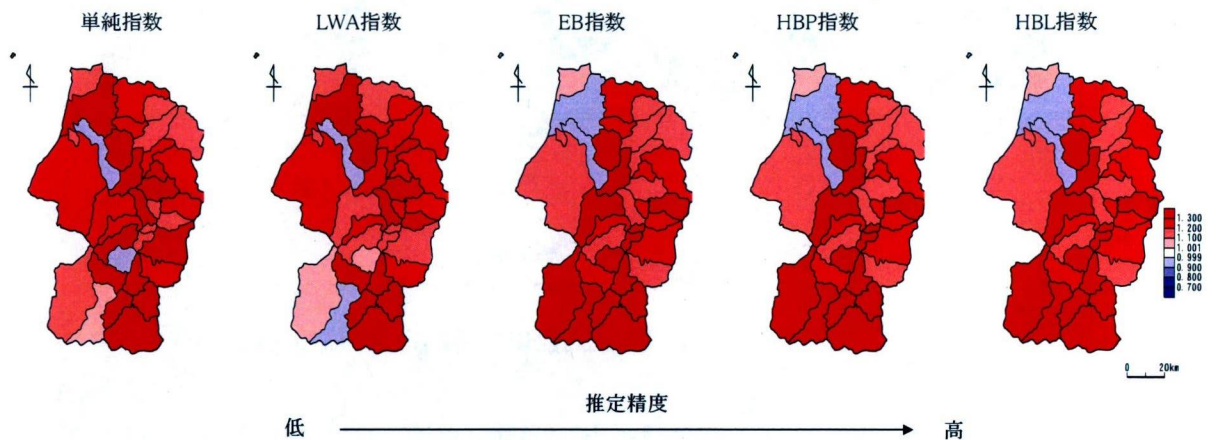


図 26 男性高血圧、受診指数の補正結果

このような期待件数が低い指数ほど大きく補正がなされることは、佐伯ほか(2008)による研究結果と同様の結果であった。佐伯ほか(2008)では、死亡率の補正について分析を行っており、結論として疾患によらず経験ベイズ法による補正を行うべきと結論付けている。ただし、この研究では SMR と経験ベイズ法の 2 つの方法(モーメント法と最尤法)による比較を行っているため、本研究とは比較しているものが異なる。本研究では、佐伯ほか(2008)の研究結果を踏まえ、補正手法はどのようなデータであっても市町村別の地域データであれば用いるべきとの前提に立ち、複数ある SAE の手法のうち、4 つの手法²⁴を選択し、そのうちいずれを用いるのが適切かを検討することに主眼を置いた。

その結果、平均二乗誤差と DIC の値から、期待件数の多い傾向にあった 3 つのデータ区分(0-39 歳高血圧、40-49 歳高血圧、50-59 歳高血圧)においては、推定精度を重視し、階層ベイズ法を選択することが適切だと考えられる。ただし、これらの区分において、経験ベイズ法を階層ベイズ法の代替手法として用いることで補正を行うことも有用だと考えられる。なぜならば、平均

²⁴ Locally Weighted Average Method、経験ベイズ法、階層ベイズ法によるポアソンガンマモデルとポアソン対数正規モデルの 4 つである。

二乗誤差の値から、階層ベイズ法による結果から著しく乖離してはいないためである。ただし、期待件数が 100-500 の場合に経験ベイズ法と階層ベイズ法による補正結果の差が大きくなる傾向がみられたため、そのようなデータを含む場合には、より慎重に手法を選択する必要がある。

また、高血圧、男性高血圧、女性高血圧、60-69 歳高血圧、70 歳以上高血圧の期待件数の大きい 5 つのデータ区分では、補正の必要性があまり高くはないため、LWA あるいは経験ベイズ法による補正手法を代替的に用いることは妥当だと考えられる。つまり、これらのデータ区分では、精度が最も良く、複雑なモデルや統計ソフトウェアを用いる階層ベイズ法を用いず、簡便な手法を用いても良いという結論に至った。上記のデータ区分においては LWA あるいは経験ベイズ法による推定結果の精度もある程度保たれているものと考えられる。

つまり、期待件数が大きなデータ区分においては、手法の簡便性を重視し、反対に期待件数の少ないデータ区分においては精度を重視した手法選択を行うべきということが確認された。以上の結果を簡略化し、データ区分ごとの補正必要度と期待件数との関係を図示したのが図 27 である。また、補正の必要度を決定する要因として、分析結果から①期待件数の小ささ、②単純指数(未補正指数)の標準偏差の大きさ、③単純指数(未補正指数)のはずれ値の有無をあげ、これらについても図 27 に示した。②と③の 2 つの要因は、0-39 歳、40-49 歳、50-59 歳、高血圧のデータ区分より観察された。

結果として、期待件数が少ない、あるいは単純指数の標準偏差が大きい、または単純指数にはずれ値がある 3 つのデータ区分(0-39 歳高血圧、40-49 歳高血圧、50-59 歳高血圧)は、補正の必要度が高く、精度を重視した階層ベイズ法による 2 つのモデルで補正を行うことが推奨される。また、高血圧のデータ区分については LWA による補正でも良いと考えられるが、期待件数のばらつきが大きいということを考慮に入れ、経験ベイズ法による補正がより適切と考えた。残りの 4 つのデータ区分(男性高血圧、60-69 歳高血圧、70 歳以上高血圧、女性高血圧)では、期待件数が大きい、単純指数の標準偏差が大きい、単純指数にはずれ値がない等の特徴があり、補正の必要性が低いと考えられる。なお、これら 4 つのデータ区分では、記述された順に補正の必要度が高いといえる。

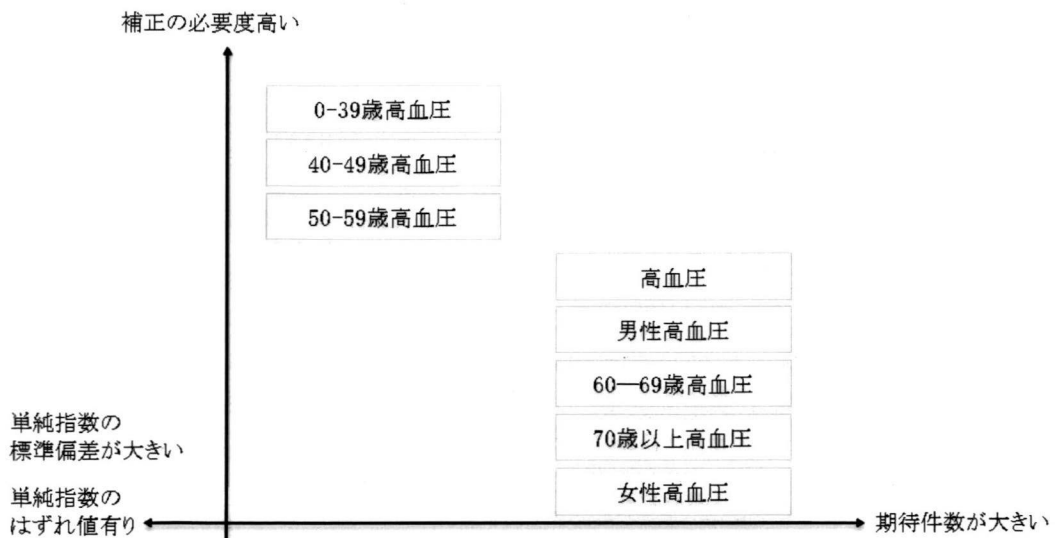


図 27 補正の必要度と 3 つの補正の決定要因

ただし、Aylin et al.(1999)が述べているように、正確さという点では経験ベイズ法は階層ベイズ法に劣ることが指摘されているため注意が必要である。本研究の分析結果においても、データ区分によっては、階層ベイズ法と経験ベイズ法による補正結果に差が生じていることも事実であり、期待件数が低いデータセットの場合には注意が必要であることもわかった。特に、期待件数が 100-500 の場合に、階層ベイズ法と経験ベイズ法の補正結果の差が大きくなる傾向が確認されたため、このようなデータ区分において常に経験ベイズ法を用いることは、精度を重視する場合には問題となる可能性がある。Devine et al.(1994)が指摘しているように、経験ベイズ法は過度に収束しすぎる(Overshrinkage)ことが欠点とされており、推定値の分布が真のリスクの分布よりも狭い傾向にあるとされている。このような欠点から、精度という観点から経験ベイズ法よりも階層ベイズ法を含むフルベイズの手法の方が優れた手法であるとの議論がなされている(C.MacNab et al.(2004))。

したがって、医療費適正化計画の策定にあたって受診指数を改良する場合には、100-500程度の期待件数を含むデータ区分の場合には、可能であれば階層ベイズ法を用いるべきであるが、代替手段として経験ベイズ法を用いることも可能だといえる。

なお、期待件数が低い場合に、補正の必要性が高いものと考えられるが、具体的に件数がいくつ以下の場合に補正が必要といえるのかについては明らかでない。本研究においては、補正差が絶対値で 0.5 以上の指数は期待件数が 500 以下となる傾向がみられたが、この点については別途詳細な研究が必要であろう。

この点について、McEwin and Elazar(2006)は、オーストラリアにおける公的な統計指標が必要となる人口規模は 10,000-200,000 人程度だと述べ、これらの人口規模において SAE の問題がしばしば生じると述べている。しかし、罹患率のあまり高くない疾患データを分析するような場合には、データ区分の人口規模が 10,000-200,000 人以上であったとしても補正の必要が生じる可能性もある。つまり、分析地域は小さくなくともデータの出現率が低い区分を分析する場合には、SAE の問題を生じる可能性があり、注意が必要である。ここでいう、データの出現率が低いデータとは、罹患率の非常に低い疾患の医療費データや、他の年齢層では罹患率が高いが分析対象となる年齢階層においては罹患率が低いような区分の医療費データを指している。

また、期待件数が低いとき、被保険者数も少ない傾向にあるといえるが、医療費データを扱う場合には、これはすべてに当てはまるとはいえない。例えば、データ区分を年齢階層別に区切った場合に、高血圧に対する罹患率があまり高くないと考えられる 0-39 歳、40-49 歳等の年齢階層では、被保険者数の人口規模が多い市町村であっても、罹患している患者は少ないために受診件数は低いという状況が生じる。反対に、疾患別・性別・年齢階層別に医療費データを分けて考えた場合、高齢になるにつれて罹患しやすい疾患であれば、人口規模がある程度小さな市町村であっても罹患している患者が多ければ SAE の問題は生じない可能性がある。

したがって、医療費データの補正を行う場合には、人口規模のみで SAE の誤差の生じ方は決定されない可能性がある。データあるいは疾病の発生頻度によって、SAE の問題の起こりやすさは異なると考えられるため、杓子定規に人口規模のみで補正の必要性を決めることはできない。分析を行う際には、対象データの期待件数がどの程度であるのかについて確認する必要があるといえよう。

そして、①人口規模の小さな地域を分析単位とする場合、②発生頻度の少ない疾患データを分