

1 Zusammenfassung der Ergebnisse der Pilotprüfung

Das Bundesamt für Soziale Sicherung (BAS) ist mit der Bitte an mich herangetreten, eine erste Einschätzung zu den Erkenntnissen aus der durchgeführten Pilotprüfung abzugeben. Diese Einschätzung ist eine Ergänzung zum Gutachten zur Risikoprüfung vom Dezember 2022. Aufgabe des Gutachtens war die Ausarbeitung einer zuverlässigen Stichproben- und Hochrechnungsmethode für die anstehenden Risikopoolprüfungen. Ziel der Risikopoolprüfungen ist eine verlässliche Abschätzung der Zuweisungsfehler in jeder Krankenkasse.

Aufgrund von Prüfkapazitätsbeschränkungen ist eine Kombination aus Voll- und Stichprobenprüfung angedacht. Das BAS hat nun eine erste Pilotprüfung in 11 Krankenkassen durchgeführt (davon 3 Vollprüfungen und 8 Stichprobenprüfungen). Hierbei wurden insgesamt 830 Hochrisikoversicherte geprüft (davon 700 in Stichprobenprüfungen und 130 in Vollprüfungen). Die Pilotprüfung wurde als einfache Zufallsstichprobe umgesetzt; hierbei stellten sich 155 Fälle als fehlerhaft heraus (18,67%). Im Mittel über alle Prüffälle sind dabei 1.524€ zu viel zugewiesen worden. Der größte Zuweisungsfehler lag bei 1.152.000€. Im Vergleich zu den anderen Zuweisungsfehlern handelt es sich hier um einen möglichen Ausreißer. Der zweitgrößte Zuweisungsfehler liegt gerade einmal bei 43.407€. Ohne diesen Datenpunkt zeigt sich ein Mittelwert von nur 137€ über alle Prüffälle. Die Diskrepanz zwischen dem Mittelwert mit und ohne Ausreißer zeigt bereits eine beträchtliche Unsicherheit auf. Diese spiegelt sich in der Standardabweichung wieder, welche mit Ausreißer bei 40.016 und ohne bei 1.671 liegt. Dies beeinflusst auch den Variationskoeffizienten (also das Verhältnis von Standardabweichung zu Mittelwert), welcher mit Ausreißer bei 26,3 und ohne bei 12,2 liegt. Im Allgemeinen gilt, je größer der Variationskoeffizient, desto größer muss die Stichprobe ausfallen, um eine gegebene Schätzgenauigkeit zu erreichen. Im Idealfall ist der Variationskoeffizient für jede Krankenkasse getrennt zu ermitteln. Mögliche Schwankungen des Variationskoeffizienten über die Krankenkassen hinweg können erst nach Vorliegen von Prüffällen aus allen Krankenkassen abschließend beurteilt werden. Auch nach der Pilotprüfung können nur unter der Annahme, dass die Variationskoeffizienten zwischen den Krankenkassen konstant sind, optimale Stichprobenumfänge festgelegt werden. Nach dem ersten Prüfzyklus sollte diese Annahme empirisch überprüft werden und/oder krankenkassenspezifische Variationskoeffizienten für den zweiten Prüfzyklus ermittelt werden.

2 Geschichtete Zufallsstichproben

Für das neue Prüfverfahren bieten sich, angelehnt an bestehende Prüfverfahren, auch geschichtete Zufallsstichproben an. Schichtungsvariablen könnten helfen, die Varianz der Schätzung zu reduzieren. Dies ist vor allem bei sehr inhomogenen Grundgesamtheiten von Vorteil und kann den damit verbundenen erhöhten Aufwand in der Umsetzung rechtfertigen (vgl. auch Hartung, 2009). Möglicherweise sind die Grundgesamtheiten jedoch in dem neuen Verfahren deutlich homogener, da diese nur Versicherte mit ohnehin erhöhtem Kostenrisiko enthalten. Vom BAS wurden die versichertenbezogenen Leistungsausgaben als mögliche Schichtungsvariable angedacht. Theoretisch sind diese fundiert, da die Zuweisungsfehler nur zwischen 0€ und den Leistungsausgaben variieren können (wenn von negativen Zuweisungsfehlern abstrahiert wird). Somit ist nicht nur ein möglicher positiver Zusammenhang im Erwartungswert zwischen Zuweisungsfehler und Leistungsausgaben denkbar, sondern auch in deren Varianz. Eine optimal geschichtete Zufallsstichprobe würde in Bereichen mit großer Fehlervarianz relativ mehr Prüffälle ziehen als in Bereichen mit kleiner Varianz. Auch eine proportionale Aufteilung des gesamten Stichprobenumfangs auf die einzelnen Schichten wäre denkbar. Wichtig ist darüber hinaus aber eine sinnvolle Aufteilung in Schichten, welche in sich möglichst ähnlich sind und sich

untereinander so weit wie möglich unterscheiden.

Erste, vom BAS erstellte Auswertungen basierend auf der Pilotprüfung liegen mir hierzu vor. Wenn man von dem möglichen Ausreißer absieht, lassen sich noch keine klaren Parameter erkennen, um ein sinnvolles Schichtungsdesign zu entwerfen. Dies ist möglicherweise dem Umstand geschuldet, dass der Stichprobenumfang in der Pilotprüfung noch verhältnismäßig klein ausgefallen ist. Gerade einmal 155 fehlerhafte Fälle stehen für die Beurteilung der möglichen Schichtungsvariable zur Verfügung. Eine sinnvolle Abgrenzung der Schichten ist durch den geringen Stichprobenumfang der Pilotprüfung deutlich erschwert. Darüber hinaus liegen gegenwärtig nur Prüffälle von 11 Krankenkassen vor, so dass eine Festlegung von optimalen Schichtungsparametern auch mit der Annahme einhergeht, dass die betrachteten Krankenkassen hinreichend aussagekräftig sind.

Ob der Aufwand einer geschichteten Zufallsstichprobe bei der Risikopoolprüfung letztendlich ausreichend Vorteile mit sich bringt, ist eine empirische Frage und bleibt auch nach der Pilotprüfung abzuwarten. Denkbar wäre als weiteres Vorgehen, den ersten Prüfzyklus als einfache Zufallsstichprobe durchzuführen und diese Ergebnisse unter anderem zu nutzen, um für zukünftige Prüfzyklen die Parameter für eine geschichtete Zufallsstichprobe zu ermitteln. In Anbetracht der bestehenden Prüfkapazitätsbeschränkungen wäre auch eine Ermittlung der Variationskoeffizienten auf Kassenartebene im ersten Prüfzyklus möglich. Dabei ist jedoch zu bedenken, dass die Schätzgenauigkeit dann nur auf dieser Ebene konstant gehalten wird und nicht notwendigerweise auch auf der Ebene der einzelnen Krankenkassen. Falls letzteres weiterhin das Ziel ist, sollten in den nachfolgenden Prüfzyklen die Variationskoeffizienten wieder krankenkassenspezifisch ermittelt werden.

3 Sicherheitsabschlag

Es liegt in der Natur der Sache, dass mit der Stichprobenprüfung eine Ungenauigkeit über die Summe der Zuweisungsfehler in der Grundgesamtheit einhergeht. Das Ziel ist, die krankenkassenspezifischen Stichprobenumfänge so zu wählen, dass die Schätzgenauigkeit der monetären Fehlerquote über alle Krankenkassen in der Stichprobenprüfung hinweg gleich ausfällt. Hierbei ist insbesondere eine genaue Ermittlung des Variationskoeffizienten notwendig. Bisher können hierzu nur die Ergebnisse der Pilotprüfung verwendet werden. Dies ist vor allem im Hinblick auf die oben erwähnten Schwankungen mit und ohne Ausreißer noch nicht zufriedenstellend (mit Ausreißer ist der Variationskoeffizient mehr als doppelt so groß). Falls man sich für den größeren Variationskoeffizienten als Basis für die Ermittlung der krankenkassenspezifischen Stichprobenumfänge entscheidet, würde das die Stichprobenumfänge bei gegebener Genauigkeit deutlich vergrößern. Falls man sich für den kleineren Koeffizienten entscheidet, aber der wahre Wert sich als größer herausstellt, kann die geplante Genauigkeit mit Hilfe der ermittelten Stichprobenumfänge nicht erreicht werden. Der erste Prüfzyklus ist daher auch nach Analyse der Pilotprüfung mit größerer Unsicherheit verbunden als nachfolgende Prüfzyklen, die auf die vorhergehenden Prüfzyklen aufbauen könnten und sollten. In bereits bestehenden Prüfverfahren wurde versucht, die Ungenauigkeit der Stichprobenschätzung durch einen Sicherheitsabschlag auszugleichen. Im Allgemeinen kann die Ungenauigkeit von Stichprobenschätzungen sowohl zu einer Überschätzung als auch zu einer Unterschätzung der tatsächlichen Summe der Zuweisungsfehler in der Grundgesamtheit führen, so dass letztendlich die Festlegung eines Sicherheitsabschlags eher eine Verhandlungsentscheidung ist und die statistische Theorie hier nicht wirklich Anhaltspunkte geben kann. Zu bedenken ist hierbei auch, wie Krankenkassen in der Vollprüfung berücksichtigt werden können. Darüber hinaus wäre es in diesem Zusammenhang auch interessant, mehr über die Ursachen der Zuweisungsfehler zu wissen – insbesondere, ob die Fehler auf systematischen Schwächen basie-

ren, welche mittelfristig beseitigt werden können, oder ob es sich um rein zufällige Fehler handelt, die auch mittelfristig unvermeidbar sind.