

SAS als Basis einer Registrierungsdatenbank: Die Alternative zu Access

Nick Martinussen Lisbeth Samsøe Schmidt
Alsak Harbo Poulsen Henrik Lind Gregersen
Joachim Schüz
Institute of Cancer Epidemiology
Danish Cancer Society
Copenhagen, Denmark

Zusammenfassung:

Epidemiologische Studien enthalten oft eine große Anzahl von Teilnehmern. Es ist wichtig, die Feldarbeit genau zu überwachen z.B. beim Verschicken von Fragebögen oder der Durchführung von Interviews, so dass man keine Teilnehmer wegen Data Management Fehlern verliert.

In Dänemark beginnen gerade zwei epidemiologische Studien zur Ursachen von Krebs, die erste mit 1500-2000 Teilnehmern und eine größere Studie mit einer möglichen Anzahl bis zu 100.000 Teilnehmern und bis zu 25 Jahren Follow-Up.

Die erste Studie mit 1500-2000 Teilnehmern, eine Fall-Kontroll-Studie zum Gebrauch vom Handys und dem Gehirntumorrisiko bei Kindern ist eine komplizierte Studie mit mehreren Möglichkeiten Teilnehmerkontakt zu verlieren, da sowohl die Kinder als auch die Eltern befragt werden müssen.

Während einer zweijährigen Zeitperiode werden alle Fälle von Gehirntumoren in der Altersgruppe 7-19 Jahre durch eine nahe Zusammenarbeit mit Pädiatrischen Kliniken identifiziert. Außerdem werden Fälle, die zwei Jahre vor dem Studienanfang diagnostiziert werden, retrospektiv eingeschlossen.

Sobald die Fälle identifiziert sind, werden zufällig Kontrollen zugewiesen und Einladungsbriefe herausgeschickt. Schon hier gibt es die erste Herausforderung, weil wenn das Kind über 18 ist, muss man nicht nur die Zustimmung zur Teilnahme bekommen sondern auch die Erlaubnis, die Eltern zu kontaktieren.

Für die zweite Studie, eine Prospektive Kohortenstudie, werden in Dänemark allein 100.000-150.000 Personen zur Studienteilnahme eingeladen. In dieser Studie gibt es nicht nur einen komplexen Kontaktvorgang sondern es ist auch eine große Datenmenge zu bewältigen, die eine Registrierungsdatenbank beeinflussen.

Unser Poster zeigt unsere Überlegungen bei der Einrichtung der entsprechenden Studierendatenbanken und deren Implementierung mit SAS.

SAS als Basis für die Registrierungsdatenbank haben wir aus mehreren Gründen gewählt. Nicht nur wegen Schnelligkeit und Zuverlässigkeit sondern auch weil wir in Zusammenarbeit mit Henrik Gregersen und seiner Entwicklungsgruppe, die Basisstruktur einer „Interviewdatenbank“ übernehmen, die wir in der erste Studie benutzen wollten.

Schlüsselwörter: Datenbank, Registrierung, Fall-Kontrolle, Kinder, Epidemiologie, Kohortenstudie, Handy, Teilnahme, Implementierung

1 Einleitung

Epidemiologische Studien enthalten oft eine große Anzahl von Teilnehmern. Es ist wichtig, die Feldarbeit genau zu überwachen z.B. beim Verschicken von Fragebögen oder der Durchführung von Interviews, so dass man keine Teilnehmer wegen Data Management Fehlern verliert.

In Dänemark beginnen gerade zwei epidemiologische Studien zu Ursachen von Krebs, die erste mit 1500-2000 Teilnehmern und eine größere Studie mit einer möglichen Anzahl bis zu 100.000 Teilnehmern und bis zu 25 Jahren Follow-Up.

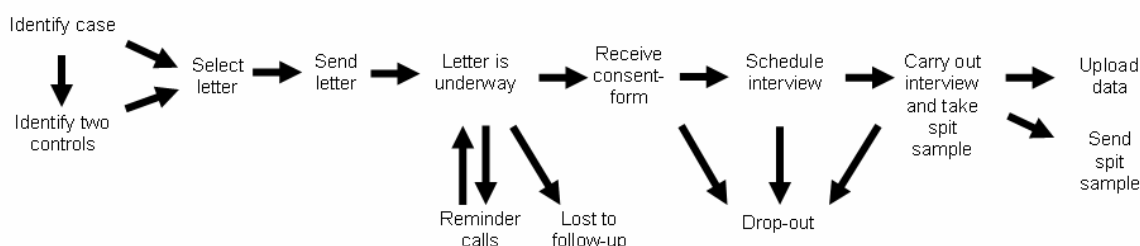
Die erste Studie mit 1500-2000 Teilnehmern - eine Fall-Kontroll-Studie zum Gebrauch vom Handys und dem Gehirntumorrisiko bei Kindern - ist eine komplizierte Studie mit mehreren Möglichkeiten Teilnehmerkontakt zu verlieren, da sowohl die Kinder als auch die Eltern befragt werden müssen.

Während einer zweijährigen Rekrutierungszeit werden alle Fälle von Gehirntumoren in der Altersgruppe 7-19 Jahre durch eine nahe Zusammenarbeit mit Pädiatrischen Kliniken identifiziert. Außerdem werden Fälle, die zwei Jahre vor dem Studienanfang diagnostiziert wurden, retrospektiv eingeschlossen.

Sobald die Fälle identifiziert sind, werden zufällig Kontrollen zugewiesen und Einladungsbriefe verschickt. Schon hier gibt es die erste Herausforderung, weil wenn das Kind über 18 ist, muss man nicht nur die Zustimmung zur Teilnahme bekommen sondern auch die Erlaubnis, die Eltern zu kontaktieren.

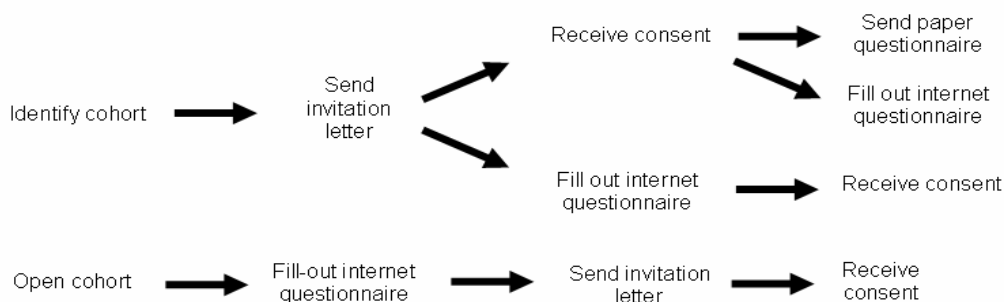
Die Kontaktvorgänge werden von uns mit Hilfe einer SAS-Datenbank überwacht.

Fig. 1. Vereinfachte Übersicht der Prozesse in der Fallkontroll-Studie



Für die zweite Studie, eine prospektive Kohortenstudie, werden alleine in Dänemark 100.000-150.000 Personen zur Studienteilnahme eingeladen (weitere Zentren bestehen in Schweden und Großbritannien). In dieser Studie gibt es nicht nur einen komplexen Kontaktvorgang sondern es ist auch eine große Datenmenge zu bewältigen, die eine Registrierungsdatenbank beeinflussen.

Fig. 2. Vereinfachte Übersicht der Prozesse in der großen Kohortenstudie



2 Ziel

Das Ziel für beide Studien ist jeweils eine Registrierungsdatenbank, mit deren Hilfe der aktuelle Status jedes Teilnehmers nachvollzogen werden kann und die selbst aktiv Vorgänge anstößt (zum Beispiel die Kontrollziehung für jeden neuen Fall oder die Meldung, dass ein Erinnerungsschreiben fällig ist), um Fehlermöglichkeiten bei der Kontaktierung von Teilnehmern zu minimieren. Die Datenbank muss transparent für jeden Nutzer sein, insbesondere für die Kohortenstudie, für die eine Laufzeit von 25 Jahren angedacht ist.

3 Strategien bei der Datenbankauswahl

SAS oder Access

Da dies die beiden Systeme waren, die im Institut verfügbar waren und eigene Erfahrungen bei der Programmierung vorlagen, stellte sich die Entscheidung zwischen Microsoft Access und SAS.

Lizenzen

Nach dem Erwerb von Access ist dieses System ohne zusätzliche Gebühren oder Lizenzen nutzbar. SAS verlangt Lizenzen um bestimmte Pakete nutzbar zu machen, was bei den meist begrenzten Finanzierungsrahmen von Studien eine Schwierigkeit darstellen kann. Für die Registrierungsdatenbanken wurde SAS/AF verwendet, weil die Lizenz bereits vorlag.

Limitationen

Oft ist Access ein erster Gedanke für Datenhaltung, aber Limitationen müssen bei der Auswahl diskutiert werden, damit sie später nicht die Prozesse behindern. So akzeptiert z.B. Access nur 255 Variablen pro Tabelle, während SAS auf 32-Bit Basis theoretisch fast 413 Millionen Variablen erlaubt.

Benutzerfreundlichkeit

Access kommt als Microsoft-Produkt dem Nutzer auf Grund von Ähnlichkeiten beim Erscheinungsbild und Nutzerführung mit Windows oder anderen Office Programmen bekannter vor. SAS nutzt ebenfalls das Windows-Erscheinungsbild für Rahmen und Eingabeboxen, mit zusätzlichen Gestaltungsmöglichkeiten kann das Erscheinungsbild weiter modifiziert werden, z.B. mit Buttons und Labels.

Integration mit Microsoft Word

Die Integration der Registrierungsdatenbank und Textverarbeitung (Word) erlaubt der Datenbank das automatische Erstellen von Kontaktbriefen. Dies ist in Access sehr einfach zu realisieren.

SAS kann über DDE (Dynamic Data Exchange) so programmiert werden, dass automatisiert erstellte Anschreiben als Wordfile generiert werden und nachbearbeitet werden können. Für die Fallkontrollstudie bestehen insgesamt neun Anschreibemöglichkeiten, z.B. je nach Alter der Kinder oder abhängig vom Vitalstatus. Diese Information steht bereits in der Registrierungsdatenbank und wird für die Brieferstellung genutzt.

4 Ergebnis

SAS als Basis für die Registrierungsdatenbank haben wir aus mehreren Gründen gewählt. Nicht nur wegen Schnelligkeit und Zuverlässigkeit sondern auch weil wir in Zusammenarbeit mit Henrik Lind Gregersen und seiner Entwicklungsgruppe, die Basisstruktur einer „Interviewdatenbank“ übernehmen, die wir in der erste Studie benutzen. Der folgende Screenshot zeigen Details zur Umsetzung. Die Registrierungsdatenbank kann als Tracking System für die Kontaktaufnahme genutzt werden und gibt eine Übersicht über die erfolgten und anstehenden Schritte.

The screenshot displays a software interface for a clinical trial registration system. On the left, a 'Today's work list' table shows tasks such as 'Carry out parent reminder call' and 'Send parent-letter' with their respective scheduled times. The central part of the screen is a 'Personal information' form with various input fields for patient and parent data, including names, addresses, and medical history. On the right, a 'Contact type' table provides a summary of contact attempts, listing the type of contact, the time it occurred, and whether it was successful.

5 Erfahrungen

Henrik Lind Gregersen, Martin Lind Gregersen, Daniel Kristoffersen und Kenneth Wi-dén haben ein computergestütztes Interviewsystem für die Fallkontrollstudie entwickelt (CAPI), dessen Struktur für die Registrierungsdatenbank genutzt wurde. Dies erlaubte eine zügige Programmierung der Funktionalität der Datenbank.

Die Adresse wird direkt aus der Datenbank in den Brief übernommen, so dass keine Fehler bei der Adressierung der Anschreiben passieren.

Die Erfahrungen unserer Projektassistenten mit der Datenbank sind sehr gut. Um Eingabefehler bei der sensiblen Studiennummer zu reduzieren, wird ein Laserstift genutzt, mit dem die Studiennummer eingelesen werden kann (auf verschickten und erhaltenen Briefen).