

Raum-zeitliche Analysealgorithmen zur Früherkennung von Gesundheitsgefahren – der SIDARTHa-Ansatz

Thomas KRAFFT, Alexandra ZIEMANN, Tim TENELSEN,
Luis GARCIA-CASTRILLO RIESGO, Matthias FISCHER, Alexander KRÄMER,
Freddy LIPPERT und Gernot VERGEINER

1 SIDARTHa-Projekt: Früherkennung von Gesundheitsgefahren in Europa

Frühwarnung vor Gesundheitsgefahren für die Bevölkerung in Europa basiert weitgehend auf traditionellen Überwachungssystemen, z. B. auf der Grundlage von bestätigten Labor-diagnosen. Diese Systeme sind nicht immer in der Lage, unbekannte Gesundheitsgefahren zu identifizieren und insbesondere aufgrund der langen Laufzeit zwischen dem Auftreten der ersten Anzeichen bis zur Meldung an die zuständige Gesundheitsbehörde ist die Frühwarnkapazität sehr eingeschränkt. Die syndrombasierte Überwachung (Syndromic Surveillance) der Inanspruchnahme von Gesundheitsversorgungsleistungen bietet die Möglichkeit, frühzeitiger mögliche Gesundheitsrisiken zu erkennen sowie entsprechend vor diesen zu warnen und ergänzt bzw. erweitert die etablierten Systeme der Gesundheitsüberwachung.

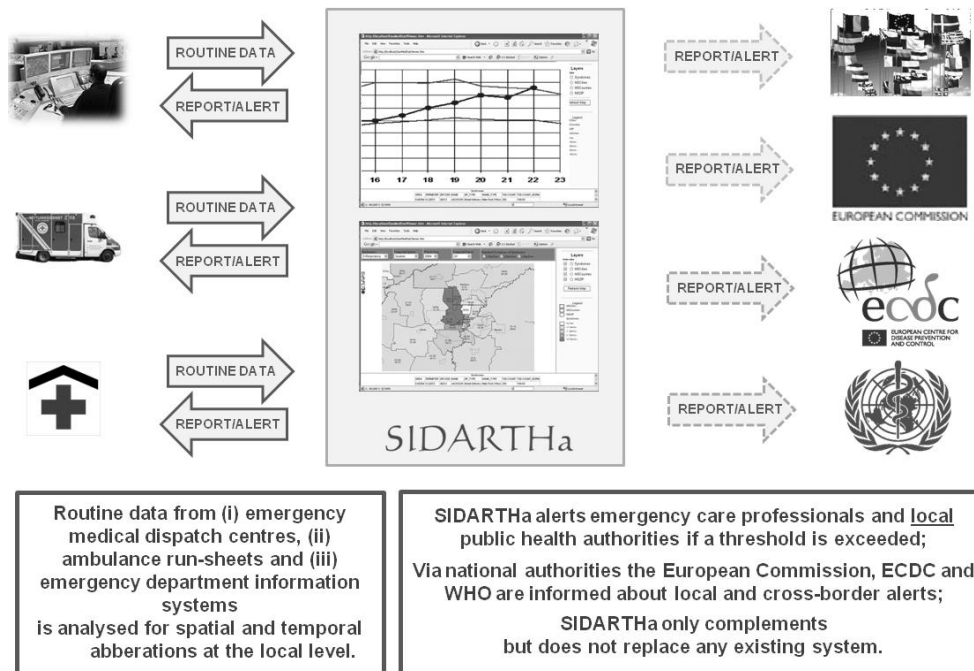


Abb. 1: Der SIDARTHa-Ansatz zur Frühwarnung vor Gesundheitsgefahren

In dem durch die Europäische Kommission geförderten Projekt SIDARTHa – European Emergency Data-based Syndromic Surveillance System for Information on, Detection and Analysis of Risks and Threats to Health (Fördernummer: 2007208; Laufzeit: Juni 2008 bis Dezember 2010, www.sidartha.eu) werden die Möglichkeiten von Syndromic Surveillance basierend auf routinemäßig erhobenen Daten des Rettungsdienstes und von Notaufnahmen der Krankenhäuser für den europäischen Kontext erstmals umfassend evaluiert. Daten der notfallmedizinischen Versorgung stellen hier eine wertvolle Informationsquelle dar, da sie europaweit, routinemäßig erhoben werden und elektronisch Informationen zu Zeit, Ort und Krankheitsbild in Echtzeit generieren (vgl. dazu u. a. KRAFFT et al. 2006, LAPOSTOLLE et al. 2005, BORK et al. 2006).

Die SIDARTHa-Projektgruppe, bestehend aus Rettungsdienstexperten, Gesundheitssystemforschern, Public Health-Vertretern und Gesundheitsgeographen aus zwölf europäischen Ländern, konzeptualisiert, entwickelt, implementiert/testet und evaluiert ein Syndromic Surveillance System, das mittels Geoinformatiktechnologie algorithmenbasiert, automatisch und kontinuierlich Routinedaten der Notfallversorgung in Echtzeit analysiert und in standardisierter Form die lokalen Gesundheitsbehörden über potenzielle übertragbare und nicht-übertragbare Gesundheitsgefahren informiert (Abbildung 1).

2 Raum-zeitliche Analysealgorithmen zur Erkennung von Gesundheitsgefahren

Die Grundlage zur Erkennung von Gesundheitsgefahren im Rahmen des SIDARTHa-Ansatzes bilden die kontinuierlichen Datenströme aus den Datenbanken der angeschlossenen Krankenhäuser und Leitstellen. Um eine potenzielle Gesundheitsgefahr (z. B. Ausbruch einer Infektionskrankheit, Kontamination von Luft oder Trinkwasser, Hitzewellen) zu erkennen, werden mit geeigneten statistischen Verfahren Basislinien und Grenzwerte der Häufigkeit einzelner Syndrome auf der Grundlage von Vergangenheitsdaten festgelegt. Das System überwacht kontinuierlich das Auftreten von Ereignissen (zeitlich) und deren Verteilung (räumlich) und vergleicht deren Häufigkeit und Verteilung mit den aus Vergangenheitsdaten vorhergesagten Werten für die jeweilige Tageskategorie und -zeit. Weichen die beobachteten Häufigkeiten und/oder Verteilungen von den vorhergesagten deutlich ab, generiert das System einen automatischen Alarm mit Informationen über die den Alarm verursachenden Entwicklungen.

Die Auswahl der Methode richtet sich dabei nach der Datenverfügbarkeit. Vornehmlich im nordamerikanischen Kontext sind eine Vielzahl von Algorithmen zur Erkennung von zeitlichen oder raumzeitlichen Clustern von Gesundheitsgefahren entwickelt worden und teils in bestehenden Surveillancesystemen im Einsatz. Für ein Gesundheitsüberwachungssystem, das auf einer Vielzahl unterschiedlicher Datenquellen mit ebenso unterschiedlichen Datenstrukturen basiert, sind einfache Modelle zur Vorhersage der Häufigkeiten besser geeignet als komplexe Modelle, da für letztere meist zusätzliche Daten benötigt werden, deren Aktualität und Verfügbarkeit nicht europaweit sichergestellt werden kann.

Zwei einfachere statistische Modelle, die sich für die Nutzung in Gesundheitsüberwachungssystemen eignen, sind das „Cumulative Sum Chart“ (CUSUM-Chart) und das „Exponentially Weighted Moving Average Chart“ (EWMA-Chart). Ursprünglich wurden sie zur statistischen Überwachung von Fertigungsprozessen entwickelt. Bei beiden Modellen

werden für den zu beobachtenden Prozess die Zustände „unter Kontrolle“ und „außer Kontrolle“ unterschieden. Zur Definition der Zustände werden eine zufällige und eine spezielle Variation der beobachteten Messwerte unterschieden. Durch ein bestimmtes Ereignis wird die spezielle Variation überschritten: es ergibt sich ein neues, unerwünschtes Prozessmittel und der Prozess gerät außer Kontrolle (vgl. LAWSON & KLEINMANN 2005, LOMBARDO & BUCKERIDGE 2007). Beide Verfahren sind leicht zu implementieren und können auf unterschiedlichen Aggregationsebenen genutzt werden. So können CUSUM- und EWMA-Charts für jedes einzelne Syndrom, so wie für das gesamte Einsatzgeschehen generiert werden. Eine räumliche Auswertung ist aber nur möglich, wenn die Charts für jede räumliche Einheit separat berechnet und ausgewertet werden.

Sind raumzeitliche aufgelöste Daten verfügbar, kommen in der Regel statistische Scanverfahren zum Einsatz, die schnell und zuverlässig auffällige Muster finden. Die SaTScan Software (KULLDORF 1997), die ursprünglich zur retrospektiven Analyse chronischer Krankheiten entwickelt wurde, kann auch im Rahmen der Syndromic Surveillance eingesetzt werden. Dazu werden dreidimensionale Zylinder mit wachsender räumlicher und zeitlicher Größe um die beobachteten Fälle gelegt. Diese Fälle können sowohl aggregiert (Bsp. pro Postleitzahlbereich), als auch als einzelne Punkte mit Koordinaten vorliegen. Der Boden des Zylinders repräsentiert die räumliche und die Höhe des Zylinders die zeitliche Komponente. Für jeden Zylinder werden die Verteilungen der beobachteten und erwarteten Fälle innerhalb und außerhalb des Kreises verglichen und ein Likelihood-Quotienten-Test durchgeführt. Die „Likelihood Funktion“ wird über alle Zylindergrößen, Zylinderhöhen und -positionen maximiert und der Zylinder mit dem höchsten Wert wird als mögliches Cluster ausgegeben. Ein Alarm wird ausgelöst, wenn das beobachtete Cluster zusätzlich signifikant anders ist als eine durch Simulation generierte, zufällige Verteilung. Durch Parametrisierung können die in SaTScan enthaltenen Scan-Verfahren auch zur prospektiven Analyse von Syndromen genutzt werden (KULLDORF 2001). Dazu ist es notwendig, durch empirische Tests eine geeignete Kombination aus minimaler und maximaler Clustergröße in sowohl räumlicher wie zeitlicher Auflösung zu bestimmen.

Obwohl CUSUM und EWMA-Charts breite Anwendung in Syndromic Surveillance Systemen gefunden haben, stellt JACKSON (2007) den Nutzen dieser Detektionsmethoden, basierend auf einem umfassenden Vergleich in Frage. Seine Kritik ist vor allem darauf gerichtet, dass die vorherrschenden Detektionsverfahren nur auf schnell steigende und mit großen Fallzahlen verbundenen Abweichungen vom Mittel reagieren, Ausbrüche, die einem anderen Verlauf folgen also nicht bemerkt werden könnten. JACKSON stellt aber auch heraus, dass die Evaluation der Detektionsalgorithmen noch ganz am Anfang steht, da die bisherigen Veröffentlichungen sich aufgrund zu unterschiedlicher Ausgangslagen oder fehlender Angaben hierzu nicht für Vergleiche eignen. Auch sind raumzeitliche Verfahren bisher noch nicht evaluiert worden. In bisherigen Studien kommt auch die Anpassung der Detektionsalgorithmen an unterschiedliche Tages- oder Saisonverläufe der Ausgangsdaten noch zu kurz (MUSCATELLO 2005, ZHU et al. 2005).

3 Der SIDARTHa-Ansatz zur raum-zeitlichen Clusteranalyse

Für das SIDARTHa Surveillancesystem wurden die verschiedenen Modelle für die Systemgestaltung und Ansätze zur Detektion evaluiert. In einem technischen Workshop der

vier Testregionen mit den technischen Entwicklungspartnern wurden, basierend auf einem detaillierten Review der angewandten Detektionsverfahren in State-of-the-Art Syndromic Surveillance Systemen weltweit, die verschiedenen Möglichkeiten diskutiert und abgewogen.

Im Rahmen von SIDARTHa sollen zunächst für vier Testregionen sowohl die zeitlichen und bei entsprechender Datenverfügbarkeit auch die raum-zeitlichen Scanverfahren implementiert und getestet werden. Die vier Testregionen sind in Dänemark (Hauptstadtregion), Deutschland (Kreis Göppingen), Österreich (Land Tirol) und Spanien (Autonome Region Kantabrien) und repräsentieren sowohl urbane, verdichtete und eher ländlich geprägte Regionen. In der Vorbereitungsphase wurden bereits lange Zeitreihen zur rettungsdienstlichen Versorgung (Tirol, Bezirk Kufstein) und zur stationären und ambulanten Notfallversorgung in Krankenhäusern (Kantabrien, Santander) ausgewertet und auf die Eignung für ein Syndromic Surveillance System getestet. Daraus wurden dann für jedes System Festlegungen für die zeitliche und räumliche Skalierung sowie zur Spezifität getroffen (vgl. Übersicht in Abbildung 2).

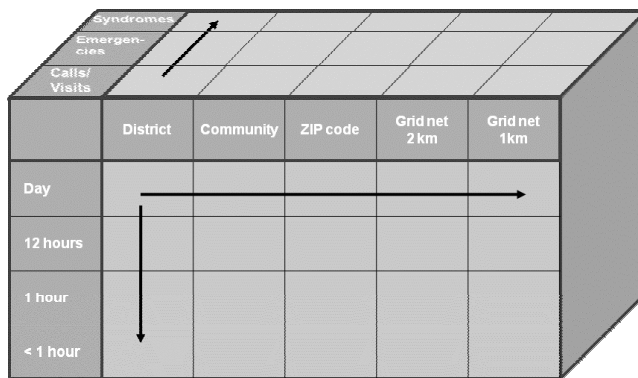


Abb. 2: Testverlauf zur Bestimmung der vorhersagbaren raumzeitlichen Basislinien und Schwellenwerte

Wegen der unterschiedlichen Bedingungen der teilnehmenden Rettungsdienste und Krankenhäuser ist die Nutzung eines universellen Verfahrens zur Bestimmung der Basislinien und Grenzwerte nicht möglich. Da die verfügbaren Daten hinsichtlich ihrer räumlichen und zeitlichen Auflösung stark variieren, muss für jedes System das geeignete Verfahren gefunden und implementiert werden (vgl. auch JACKSON 2007). Simulationen sind dabei die beste Herangehensweise, um das jeweils geeignete Verfahren sowie die kleinste zeitliche und räumliche Einheit zu bestimmen, für die eine Vorhersage statistisch valide erfolgen kann. Für die in Rettungsleitstellen erhobenen Punktdaten (Einsatzort) werden Scanverfahren eingesetzt, wohingegen die auf Postleitzahlebene erhobenen Patientenwohnorte in den Krankenhäusern primär mit den Methoden der Zeitreihenanalyse bearbeitet werden. Das Surveillancesystem kann so auf lokal spezifische Ausprägungen hinsichtlich der betrachteten Syndrome, aber auch hinsichtlich der Technik und Datenverfügbarkeit ausgerichtet werden.

Im lokal-regionalen Systemansatz von SIDARTHa werden die Daten vor Ort in den beteiligten Notfallversorgungsinstitutionen über eine SIDARTHa Softwareapplikation ausgewertet. Die Applikation (Abbildung 3) übernimmt automatisch die Datenbereinigung und überführt die Daten in zwei Auswertestränge: 1. zur Syndromgenerierung (den SIDARTHa-Festlegungen folgend) und 2. zur Aktualisierung der raumzeitlichen Basislinien- und Schwellenwertberechnung. Der aktuelle Verlauf des Notfallgeschehens (nach Syndromen sortiert) wird mit den lokal angepassten Schwellenwerten (für die einzelnen Syndrome) verglichen, sodass automatisch ein Alarm ausgelöst werden kann, wenn der Schwellenwert überschritten wurde. Es verlassen also keine Rohdaten die Notfallinstitution, sodass die Datenschutzbestimmungen eingehalten werden. Der Alarm wird zum einen an die Notfallinstitution selbst sowie an den lokalen/regionalen Gesundheitsdienst weitergeleitet, der den Alarm zurückverfolgt und über die etablierten nationalen Berichtswege auf die regionale/nationale und die europäische Ebene weiterleiten kann. SIDARTHa wird also auf lokaler/regionaler Ebene nur die bestehenden Surveillancestrukturen unterstützen und verbessern, aber keine Parallelstrukturen für die Berichterstattung schaffen. Eine Weiterleitung auf überregionale Ebene (zeitlich und räumlich verortet) kann eine raumzeitliche Betrachtung einer Gesundheitsgefahr über Grenzen hinweg ermöglichen (SIDARTHa European Health Surveillance System). Diese nur bestimmten Nutzergruppen zugängliche Funktion wird z.Zt. mit den bestehenden nationalen und europäischen Systemen abgeglichen, um die beste Möglichkeit einer Integration zu finden. So könnte beispielsweise das europäische Frühwarnsystem MedISys (Medical Intelligence System – ein automatisches System der EU, das bestimmte medizinische Webseiten und Nachrichtenwebseiten auf Hinweise zu Gesundheitsgefahren scannt) das SIDARTHa European Health Surveillance System scannen und in seine Berichtsstrukturen einfügen. Das SIDARTHa European Health Surveillance System soll aber v.a. den in der Notfallversorgung verantwortlich Tätigen ermöglichen, sich mit ihren Kollegen in andern Notfalleinrichtungen und den lokalen Gesundheitsbehörden über eine Gesundheitsgefahr in ihrer Region auszutauschen und ihre Maßnahmen zu koordinieren (Emergency Professionals Risk Communication Platform).

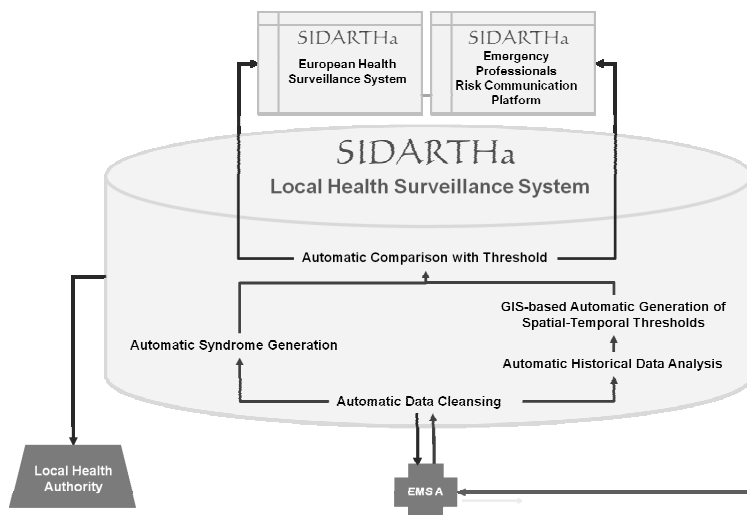


Abb. 3: SIDARTHa – Design des lokalen Gesundheitsüberwachungssystems

4 **Ausblick:** **Aufbau und Einsatz des SIDARTHa Surveillance-Systems**

Mit der Fokussierung des SIDARTHa-Ansatzes auf die lokale/regionale Ebene wird die Verbreitung des Syndromic Surveillance Systems nach der Projektlaufzeit über die Projektpartner hinaus unterstützt.

Die identifizierten raum-zeitlichen Analysealgorithmen werden für die vier Testgebiete in das jeweilige lokale SIDARTHa Surveillance System in automatisierter Form integriert. Die Implementierung der Systeme wird bis Ende November 2009 abgeschlossen sein. Bis Ende 2010 werden die Systeme in Praxisläufen eingehend getestet, von den Nutzern evaluiert und – soweit notwendig – weiter angepasst.

Literatur

- BORK, K. H., KLEIN, B. M., MOLBAK, K., TRAUTNER, S., PEDERSEN, U. B. & HEEGAARD, E. (2006): Surveillance of ambulance dispatch data as a tool for early warning. In: *Eurosurveillance*, 11 (10-12), S. 229-233.
- JACKSON, M. L., BAER, A., PAINTER, I. & DUCHIN, J. (2007): „A simulation study comparing aberration detection algorithms for syndromic surveillance.“ *BMC Medical Informatics and Decision Making*, 7. S. 6.
- KRAFFT, T., GARCIA-CASTRILLO RIESGO, L., FISCHER, M., LIPPERT, F., OVERTON, J. & ROBERTSON-STEEL, I. (2006): *Health Monitoring & Benchmarking of European EMS Systems: Components, Indicators, Recommendations*. Köln.
- KULLDORFF, M. (1997): A spatial scan statistic. In: *Communications in Statistics: Theory and Methods*, 26. S. 1481-1496.
- KULLDORFF, M. (2001): Prospective time-periodic geographical disease surveillance using a scan statistic. In: *Journal of the Royal Statistical Society*, A164. S. 61-72.
- LAPOSTOLLE, F., FLEURY, M., CROCHETON, N., GALINSKI, M., CUPA, M., LAPANDRY, C. & ADNET, F. (2005): Early indicators of a health crisis: Examination of the records of an emergency medical services call center during the August 2003 heat wave. In: *La Presse Medicale*, 34. S. 199-202
- LAWSON, A. B. & KLEINMANN, K. (Eds.) (2005): *Spatial and Syndromic Surveillance for Public Health*. Chichester.
- LOMBARDO, J. S. & BUCKERIDGE, D. L. (Eds.) (2007): *Disease surveillance: a public health informatics approach*. Chichester.
- MUSCATELLO, D. J., CHURCHES, T., KALDOR, J., ZHENG, W., CHIU, C., CORRELL, P. & JORM, L. (2005): An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. In: *BMC Public Health*, 5, S. 141.
- ZHU, Y., WANG, W., ATRUBIN, D. & WU, Y. (2005): Initial evaluation of the early aberration reporting system--Florida. In: *Morbidity and Mortality Weekly Report*, 54 (Suppl), S. 123-130.