

Hautarzt 2022 · 73:53–60  
<https://doi.org/10.1007/s00105-021-04918-x>  
 Angenommen: 22. Oktober 2021  
 Online publiziert: 23. November 2021  
 © The Author(s), under exclusive licence to  
 Springer Medizin Verlag GmbH, ein Teil von  
 Springer Nature 2021



# Websearch-Daten als Gesundheitsdaten?

## Geografische Unterschiede, zeitliche Trends und Interessenschwerpunkte von Internetsuchmaschinenanfragen in Deutschland

S. Ziehfreund<sup>1,2</sup> · L. Tizek<sup>1</sup> · A. Zink<sup>1,3</sup>

<sup>1</sup> Technische Universität München, Fakultät für Medizin, Klinik und Poliklinik für Dermatologie und Allergologie am Biederstein, München, Deutschland

<sup>2</sup> Technische Universität München, Fakultät für Medizin, Institut für Allgemeinmedizin und Versorgungsforschung, München, Deutschland

<sup>3</sup> Division of Dermatology and Venereology, Department of Medicine Solna, Karolinska Institute, Stockholm, Schweden

### Zusammenfassung

**Hintergrund:** Die eingeschränkt zeitnahe, kostengünstige und regionale Verfügbarkeit von Daten im Gesundheitswesen gilt als einer der limitierenden Faktoren für zeitgerechte Analysen im Bereich der Versorgungsforschung und damit für die Gesundheitsversorgung der allgemeinen Bevölkerung. Deshalb geraten zunehmend Internetsuchmaschinenanalysen in den Fokus.

**Fragestellung:** Welchen Beitrag können Daten zum Internetsuchvolumen zu verschiedenen Erkrankungen in der Gesundheitsversorgung der deutschen Bevölkerung leisten? Wo liegt das Potenzial, und wo gibt es Grenzen?

**Material und Methoden:** Es erfolgten die Aufbereitung der aktuellen Literatur sowie eine selektive Übersicht der in den letzten 4 Jahren veröffentlichten Suchmaschinenanalysen zu dermatologischen, allergologischen und infektiologischen Erkrankungen in Deutschland.

**Ergebnisse:** Durch die Analyse von Suchmaschinendaten konnten zeitliche Entwicklungen wie Saisonalität, Monate mit dem höchsten Suchinteresse und geografische Unterschiede auf nationaler, Bundesländer- und Städteebene abgebildet und Interessenschwerpunkte bezüglich eines die Krankheit betreffenden Themas (z. B. Hautkrebs am Auge oder analer Juckreiz) aufgezeigt werden. Darüber hinaus fanden manche Studien einen Zusammenhang zwischen dem Suchvolumen und externen Faktoren (z. B. Temperatur, ärztliche Versorgungsstruktur) sowie zu registrierten Fällen (z. B. Hautkrebs, Borreliose).

**Schlussfolgerung:** Internetsuchmaschinendaten liefern als nahezu in Echtzeit verfügbare Datenquelle unter Berücksichtigung der aufgezeigten Fallstricke ein räumlich-zeitliches Abbild hinsichtlich der Bedürfnisse der internetnutzenden Bevölkerung. Sie können besonders nützlich in Situationen sein, in denen traditionelle Gesundheitsdaten begrenzt oder nicht vorhanden sind.

### Schlüsselwörter

Geografie · Medical needs · Google · Dermatologie · Infektiologie

### Zusatzmaterial online

Die Online-Version dieses Beitrags (<https://doi.org/10.1007/s00105-021-04918-x>) enthält eine zusätzliche Tabelle. Beitrag und Zusatzmaterial stehen Ihnen im elektronischen Volltextarchiv auf <https://www.springermedizin.de/der-hautarzt> zur Verfügung. Sie finden das Zusatzmaterial am Beitragsende unter „Supplementary Information“.

### Additional Information

Die Tabelle (Online-Tab. 1) ist aus Platzgründen nur elektronisch als Supplementary Information verfügbar.



QR-Code scannen & Beitrag online lesen

### Daten im Gesundheitswesen

Medizinische Daten, Public-Health-Daten, Versicherungsdaten, Forschungsdaten und durch das Individuum generierte Daten – die Datenlandschaft im Ge-

sundheitswesen ist sehr vielschichtig. Allerdings kann zumeist noch nicht das volle Potenzial ausgeschöpft werden, denn die meisten genannten Daten sind aufgrund von Übermittlungszeiten sowie administrativen und qualitätssichernden

Prozessen sehr zeitversetzt zugänglich und nicht unmittelbar für die Versorgungsplanung und -steuerung verwertbar [8]. In Notfallsituationen wie der COVID-19-Pandemie oder regionalen Infektionsausbrüchen (z.B. Skabies), in denen schnelle, datenbasierte Entscheidungen notwendig sind, kann dieser Zeitverzug besonders problematisch sein. Andersherum sind bestimmte Daten zwar nahezu in Echtzeit, aber keineswegs flächendeckend verfügbar, z.B. Echtzeitdaten zur Pollenkonzentration [18]. Gleichzeitig kommt es durch spezielle rechtliche und ethische Rahmenbedingungen in der Medizin und dem Gesundheitswesen, insbesondere durch die Tatsache, dass die Auswertung von sensiblen und damit besonders schützenswerten Daten oftmals im Spannungsfeld konkurrierender Interessen oder Verordnungen steht, hierzulande noch zu starken Einschränkungen [32]. Ein weiteres Problem ist, dass es für zahlreiche medizinische Beschwerden und Krankheitsbilder schwierig ist, die Krankheitslast für Deutschland abzuschätzen, da sie nicht oder nur eingeschränkt meldepflichtig sind (z.B. Borreliose, Chlamydien) [13] oder nicht in Registern erfasst werden (z.B. weißer Hautkrebs) [37]. Darüber hinaus gibt es immer noch tabuisierte Krankheiten, über die viele Betroffene nicht mit dem medizinischen Fachpersonal sprechen [1], oder Betroffene, die erst gar keine medizinischen Leistungen in Anspruch nehmen [23, 33]. Folglich kommt es zu schwerwiegenden Einschränkungen für eine optimale spatiotemporale Allokation von Ressourcen in der medizinischen Versorgung und dem Gesundheitswesen wie bei der Implementierung zielgerichteter Präventionskampagnen und Interventionen, aber auch der personenzentrierten Gesundheitsversorgung.

### Digitalisierung im Gesundheitsbereich

Das Sekundärdatenkonzept schließt grundsätzlich keine Datenquelle aus. Vielmehr kann es als Anforderung angenommen werden, alle verfügbaren Daten auf ihre Anwendbarkeit „als Basis für Auswertungen im Sinne der wissenschaftlichen Fragestellung“ zu prüfen [28]. Angetrieben durch die Digitalisierung des

Gesundheitssektors, ergaben sich eine beinahe unbegrenzt vielfältige Datenbasis und die Möglichkeit zu innovativen Datenanalysen: Gesundheits-Apps und Wearables, die das Gesundheitsverhalten messen, das Krankheitsgeschehen dokumentieren und Real-time-Aktivitätsprofile erstellen [17, 18, 27], gesundheitsbezogene Informationen von auf partizipativen Online-Plattformen oder sozialen Medien geteilten Inhalten [5] und Suchanfragen über Internetsuchmaschinen; alles Daten(quellen), die als Basis zur Sekundärdatenanalyse herangezogen werden können.

Die Analyse der Internetsuchmaschinenanfragen erweist sich dabei als ein vielversprechender Ansatz zur Untersuchung des Interesses hinsichtlich der die Gesundheit und Krankheit betreffenden Themen innerhalb der Bevölkerung. Denn nicht nur Betroffene und Patient\*innen informieren sich im Internet über Symptome, Ursachen, Einflussfaktoren, und Behandlungsmöglichkeiten ihrer Beschwerden, sondern auch deren Familienangehörige und Freund\*innen. Zudem suchen Leute online nach Informationen zur Prävention, Gesundheitsförderung, zum Gesundheitsverhalten, zu Krankheiten und zu spezifischen medizinischen Belangen (z.B. Substanzabhängigkeit, finanzielle Unterstützung bei schwerwiegenden Erkrankungen) [6]. Dies eröffnet die Möglichkeit, das Interesse der Bevölkerung außerhalb des klassischen medizinischen Settings und über Routinedaten hinaus zu identifizieren.

### Analyse von Internetsuchmaschinen Daten in Deutschland

In Deutschland nutzten zwischen 2018 und 2021 bis zu 95 % der über 13-jährigen Bevölkerung das Internet, und mit ca. 90 % ist Google die am häufigsten verwendete Suchmaschine [12]. So kann durch Google-Suchmaschinenanalysen unter Zuhilfenahme von den Tools *Google Ads Keyword Planner* und *Google Trends* das Interesse der internetnutzenden Bevölkerung in Deutschland abgebildet werden.

*Google Ads Keyword Planner* und *Google Trends* sind in erster Linie dazu gedacht, um Marketingkampagnen zu optimieren. Aufgrund der Möglich-

keit, das Interesse der Bevölkerung an bestimmten Themen widerzuspiegeln, werden die Tools immer häufiger auch zur Beantwortung von wissenschaftlichen Fragestellungen herangezogen [20, 25, 29].

*Google Ads Keyword Planner* identifiziert, basierend auf einem für die Forschungsfrage ausgewählten Keyword (z.B. Psoriasis) oder Keywordcluster (z.B. Basaliom, weißer Hautkrebs, Basalkarzinom), eine Liste mit assoziierten Suchbegriffen (= einzelne Wörter oder Phrasen, die von den Nutzer\*innen in die Suchmaske von Google eingegeben werden) inklusive deren geschätztem monatlichem Suchvolumen (= Suchanfragen, die pro Monat bei Google generiert werden). Diese Daten sind vom Tag der Abfrage für die letzten 4 Jahre verfügbar.

*Google Trends* dagegen liefert keine Informationen bezüglich verwendeter Suchbegriffe, sondern das relative Suchvolumen (RSV), das zwischen 0 und 100 schwankt (100=höchstes Interesse im betrachteten Zeitraum und der betrachteten Region, 0=geringstes Interesse). Das RSV ist angepasst an die Anzahl an Nutzer\*innen in einer Region und einen bestimmten Zeitraum und misst demnach die Intensität der Suchanfragen. Die Daten sind seit 2004 für jeden Tag bis wenige Tage vor dem Zugriffsdatum verfügbar.

Mittels der so identifizierten Daten lassen sich für spezifische Erkrankungen (z.B. Psoriasis, atopische Dermatitis [AD]), aber auch für übergeordnete Krankheitsbilder (z.B. Hauterkrankungen, sexuell übertragbare Erkrankungen [STI]) die folgenden Fragestellungen untersuchen: Für was interessiert sich die Bevölkerung und wo liegen sog. „unmet medical needs“? Wie verändert sich das Interesse im Laufe der Zeit? Gibt es regionale Unterschiede im Suchverhalten? Lassen sich Zusammenhänge mit externen Faktoren wie klimatischen Bedingungen oder der Versorgungsstruktur identifizieren? Lassen sich Rückschlüsse auf die Krankheitslast ziehen?

### Interessenschwerpunkte und „medical needs“

Unabhängig von der untersuchten Thematik lassen sich anhand der von *Google Ads Keyword Planner* identifizierten, re-

levanten Suchbegriffe und dem dazugehörigen Suchvolumen Interessenschwerpunkte der Bevölkerung feststellen (Online-Tab. 1; [11, 14–16, 20, 22, 25, 29–31, 34, 38, 39]). So zeigte sich beispielsweise bei Analysen zu *Hautkrebs*, gemessen an der Anzahl an Suchbegriffen, ein großes Interesse an der Identifikation von Hautkrebs und an auftretenden Symptomen [25, 31]. Dies deutet darauf hin, dass Menschen sich online informieren, noch bevor sie medizinisches Fachpersonal aufsuchen [9]. Bezüglich *Geschlechtskrankheiten* war ein beachtlicher Teil des identifizierten Suchvolumens zu Chlamydien [20]. Wie von den Autoren beschrieben, könnte dieses Interesse auf das durch die Krankenkasse gezahlte Chlamydienscreening für Frauen bis 24 Jahre zurückzuführen sein, bei dem die meisten Frauen mit der Erkrankung in Berührung kommen. Zudem stellen Chlamydien die häufigste bakterielle STI in Deutschland dar, was vermuten lässt, dass das Suchvolumen Rückschlüsse auf Krankheitshäufigkeiten zulässt [20].

### » Menschen informieren sich online, noch bevor sie medizinisches Fachpersonal aufsuchen

Bei Analysen zu mit Juckreiz assoziierten Erkrankungen (AD, Psoriasis, Skabies sowie Pruritus) ließ sich ein Großteil der Keywords bestimmten Lokalisationen zuordnen [14, 16, 29, 30, 38, 39]. Diesbezüglich berichteten Tizek et al. [30] allerdings eine teilweise konträre Verteilung der Suchanfragen verglichen mit der in der Literatur beschriebenen klinischen Verteilung der Lokalisationen der AD. So waren Suchanfragen zu den unteren Extremitäten verhältnismäßig seltener, als in der Literatur berichtet, und der Anteil der Suchanfragen zu Lokalisationen im Gesicht doppelt so hoch beschrieben. Dies begründeten die Autoren mit einem höheren Leidensdruck und dem Wunsch nach weiteren Informationen, wenn das Gesicht betroffen ist, und verwiesen darauf, bei der medizinischen Versorgung besonderes darauf zu achten [30]. Solche medizinischen Bedürfnisse konnten z.B. auch im Rahmen der Analysen zu *Hautkrebs* bezüglich Melanom im Auge, zu *Sarkoidose* bezüglich der Neurosarkoidose und zu *Pruritus* in Bezug auf analen Juckreiz [25, 38, 39] iden-

tifiziert werden. Dies verdeutlicht, dass die Analyse von Suchmaschinenanfragen die Identifikation von essenziellen Faktoren, die bei den Suchenden von besonderem Interesse sind oder die im gesellschaftlichen Diskurs nur unterschwellig präsent sind, ermöglicht.

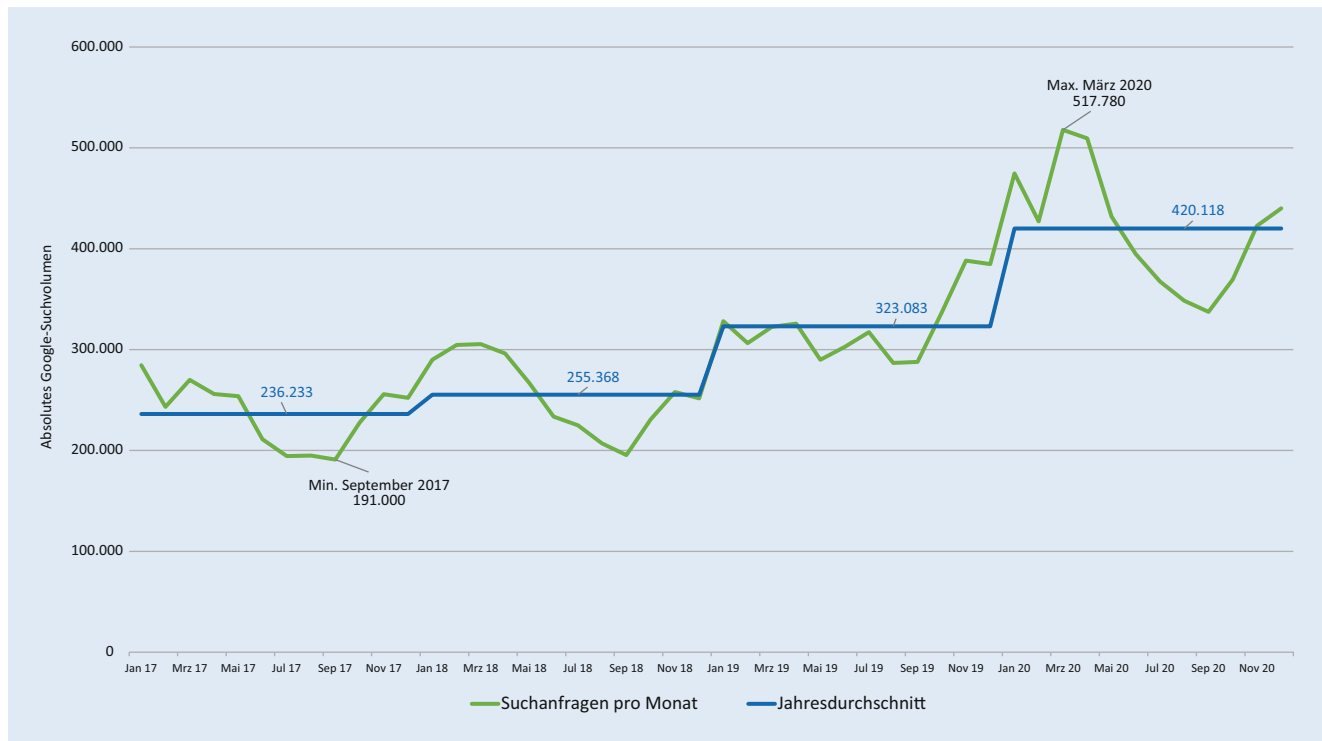
### Zeitliche Trends und Zusammenhang mit externen Faktoren

Das Suchvolumen kann auch im zeitlichen Verlauf dargestellt werden, sodass Trends während des Untersuchungszeitraums aufgezeigt werden können. Ein Anstieg im Suchvolumen [11, 14, 25, 30] kann durch die zunehmende Nutzung von Suchmaschinen als Quelle für Gesundheitsinformationen erklärt werden, wie beispielsweise bei Skabies, was unter anderem durch an der Zunahme der diagnostizierten Fällen im betrachteten Zeitraum liegen kann [14]. Weitere Untersuchungen demonstrierten die Fähigkeiten der Daten zur Darstellung der jeweiligen Krankheitslast im Zeitverlauf [15, 22, 31]. Etwa korrelierte das Suchvolumen zu *Hautkrebs* positiv mit den entsprechenden Krebsregisterdaten [31], das Suchvolumen zu *Borreliose* mit den registrierten Fällen [22] und das Suchvolumen zu *Bindehautentzündung* mit den in einer Klinik diagnostizierten Fällen [15]. Kammrath Betancor et al. [15] berichten zudem in ihrer Arbeit zu Bindehautentzündungen, dass trotz des zeitlichen Zusammenhangs die Notfallpraxis der Häufigkeit hinterherhinkte, was damit erklärt wurde, dass lediglich ein kleinerer Teil der Betroffenen mit schwereren oder langwierigeren Symptomen mit größerer Wahrscheinlichkeit zeitverzögert tatsächlich eine Augenärztin oder einen Augenarzt aufsucht.

Es können auch saisonale Trends im Suchverhalten analysiert und identifiziert werden, die sich mit zyklischen externen Faktoren in Verbindung bringen lassen [11, 15, 16, 22, 25, 29, 31, 38, 39]. Exemplarisch können hier *Hautkrebs* und eine erhöhte UV-Strahlung in den Frühjahr- und Sommermonaten [25, 31] sowie *Pruritus* und *Xerosis cutis* in den Wintermonaten genannt werden [29, 38, 39]. Folglich konnte in Analysen zu *Pruritus* und *AD* gezeigt werden, dass ein Anstieg der Temperatur und der Sonnenstunden mit einem Rück-

Hier steht eine Anzeige.

 Springer



**Abb. 1** ▲ Zeitlicher Trend des Google-Suchvolumens zu atopischer Dermatitis von Januar 2017 bis Dezember 2020

gang im Suchvolumen einherging [16, 22, 29]. Im Gegensatz dazu konnte ein Anstieg im Suchvolumen zu *Borreliose* mit einem Anstieg der Temperatur im späten Frühjahr, wenn Zecken aktiv werden, in Verbindung gebracht werden [22]. Andere Untersuchungen, beispielsweise zum Thema *Geschlechtskrankheiten*, zeigten hingegen keine wesentlichen saisonalen Trends, was laut Autoren darauf schließen lässt, dass das Interesse an Geschlechtskrankheiten unabhängig von periodisch externen Faktoren wie Klima ist [20]. Jedoch zeichneten sich in manchen Monaten deutliche Peaks im Suchvolumen zu Geschlechtskrankheiten ab, die womöglich auf eine verstärkte Medienberichterstattung zurückzuführen waren [20].

### » Berichterstattung durch die Medien hat einen wesentlichen Einfluss auf das Suchverhalten

Dass die Berichterstattung durch die Medien einen wesentlichen Einfluss auf das Suchverhalten hat, konnte in einer *Google Trends* Analyse zu COVID-19 dargestellt werden [24]. In der besagten Analyse konnten Schwankungen im Suchvolumen eher mit der Medienberichterstattung, z. B. ers-

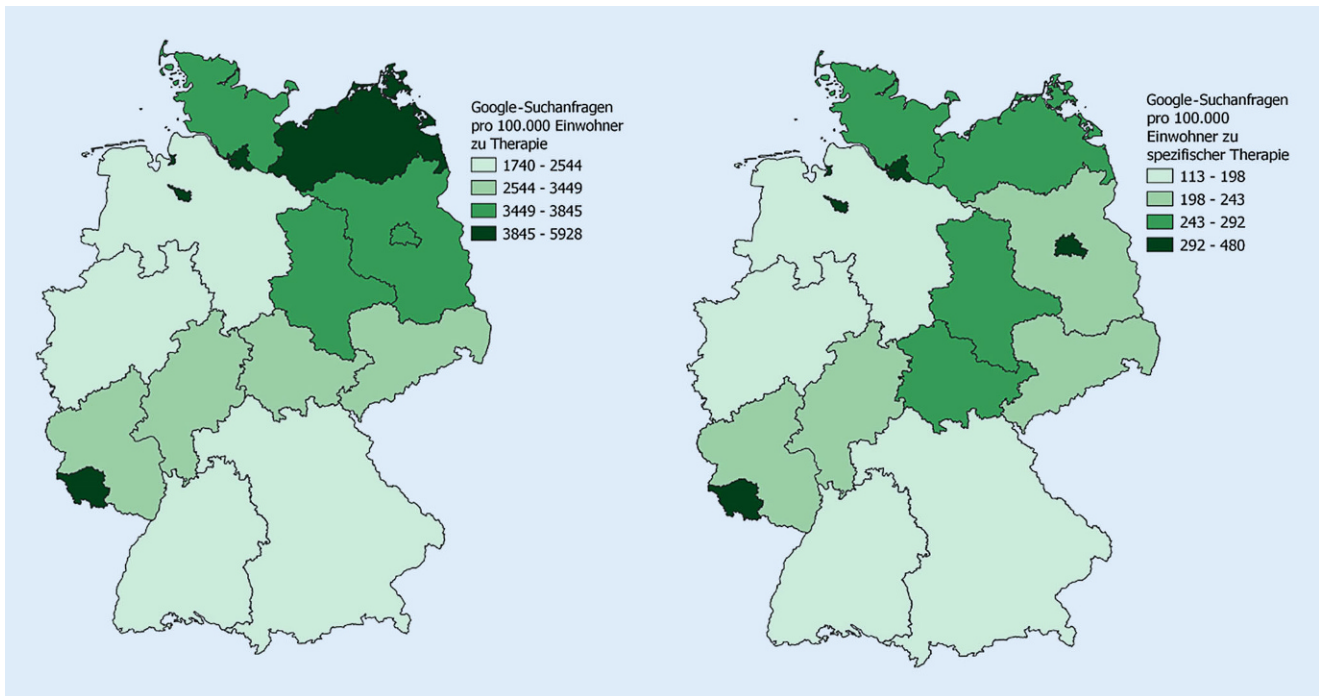
ter COVID-19-Fall in einem Bundesland, als mit epidemiologischen Daten in Verbindung gebracht werden [24]. Zudem können bedeutende Geschehnisse für die Bevölkerung wie die COVID-19-Pandemie augenscheinlich Auswirkungen auf das Suchverhalten bezüglich anderer Erkrankungen haben. So war auffällig, dass das ohnehin große Interesse, sich online über AD zu informieren, durch die COVID-19-Pandemie noch einmal deutlich größer geworden ist (Abb. 1). Dies führten die Autoren darauf zurück, dass zur Pandemiezeit weniger Menschen dermatologische Leistungen in Anspruch genommen haben [35].

In der Gesamtheit betrachtet, lassen sich aus den zeitlichen Verläufen Informationen ableiten, die von grundlegender Bedeutung sind, um unter anderem evidenzbasierte Informationen genau in dem Moment bereitzustellen, wenn Informationsbedürftige danach suchen. Diese könnten beispielsweise über Suchmaschinen für die Zielgruppe intelligent platziert werden oder über häufig genutzte Online-Plattformen zur Verfügung gestellt werden.

### Regionale Unterschiede

Das Suchvolumen lässt sich nicht nur auf nationaler Ebene beleuchten. Auch eine differenzierte Betrachtung auf Bundesländer- und Städteebene sowie zwischen ländlichen und urbanen Regionen ist möglich. Für diesen Zweck wird das Suchvolumen auf die jeweilige Einwohnerzahl (Suchvolumen/100.000 bzw. 1000 Einwohner) berechnet, um die Daten vergleichbar zu machen. Damit kann das Suchvolumen zu einem bestimmten Thema auf räumliche Gemeinsamkeiten und Unterschiede hin untersucht werden, um mögliche Interessenschwerpunkte für die jeweiligen Gebiete zu identifizieren.

Was dies betrifft, zeigte beispielsweise eine Analyse des Internetsuchvolumens zu *Skabies*, dass die Suchanfragen für den Stadtstaat Hamburg und das sehr dicht besiedelte Bundesland Nordrhein-Westfalen mehr als doppelt so hoch waren als die für das ländlich geprägte Bayern, weshalb eine positive Korrelation zwischen dem Suchvolumen und der Bevölkerungsdichte festgestellt wurde [14]. Eine mögliche Erklärung war laut Autoren die Tatsache, dass Skabies in Regionen, in denen die Menschen näher beieinander leben, schneller über-



**Abb. 2** ▲ Google-Suchmaschinenanalyse zu Psoriasis. Vergleich von mit „Behandlung“ und „spezifischer Therapie“ in Verbindung stehendem Suchvolumen in den 16 Bundesländern

tragen werden kann, was das öffentliche Interesse an dem Thema erhöht. Außerdem, so die Autoren, könnten das niedrigere Durchschnittsalter und die damit einhergehend höhere Internetnutzung in den urbanen Regionen einen Einfluss auf das Suchvolumen haben [14]. Neben dem Durchschnittsalter kann auch der Frauenanteil einen Einfluss auf das Suchverhalten haben. Dies schlussfolgerten Mick et al. [16] als Folge des unterschiedlich hohen Suchvolumens zu AD zwischen deutschen Städten, wobei ein höheres Suchvolumen mit einem niedrigeren Durchschnittsalter und einem höheren Frauenanteil in Zusammenhang gebracht werden konnte [16].

Regionale Unterschiede im absoluten Suchvolumen auf Städteebene konnten unter anderem in Analysen zu *Pruritus* [29] und *Hautkrebs* [31] herausgefunden werden. So gab es in Nürnberg und Köln im Jahr 2014 besonderes viele Melanomsuchanfragen/100.000 Einwohner [31], wobei die Krebsregisterdaten zur altersstandardisierten Melanominzidenz ebenfalls in diesen Städten vergleichsweise hoch waren. Angelehnt daran, konnten Scheerer et al. [22] eine (sehr) hohe Korrelation zwischen dem Suchvolumen zu

*Borreliose* und registrierten Borreliosefällen auf Bundeslandebene identifizieren.

Außerdem konnten Tizek et al. [29] zwischen dem Suchvolumen zu *Pruritus* und der hausärztlichen, aber besonders der dermatologischen Versorgung in den untersuchten Städten eine hohe Korrelation herausfinden. Ebenso wiesen Suchmaschinendaten zu *Psoriasis* einen ähnlichen Effekt auf [34], wobei ein bedeutender Unterschied zwischen den Bundesländern in Bezug auf das behandlungsbezogene Suchvolumen entdeckt wurde (■ Abb. 2). So korrelierte die Verschreibungsrate von Biologika positiv mit dem Suchvolumen zu mit „Behandlung“ und „spezifischer Therapie“ in Verbindung stehenden Suchbegriffen (z.B. „Psoriasis Behandlung“ bzw. „Biologika Psoriasis“). Laut Autoren verdeutlicht dies die unterschiedlichen Therapieangebote in den einzelnen Bundesländern, was besonders auffällig beim Ost-West-Gefälle war (■ Abb. 2). Dementsprechend schlussfolgerten die Autoren, dass der Zugang zur medizinischen Versorgung bzw. Therapie das Suchverhalten der Bevölkerung zu beeinflussen scheint [34]. Im Rahmen der Analyse von AD-Daten konnte festgestellt werden, dass Schleswig-Holstein den höchsten Anteil an Suchbegriffen be-

züglich Biologika hatte, wohingegen Bremen einen sehr hohen Anteil an Suchanfragen zu Hausmitteln verzeichnete. Letzteres könnte, laut Autoren, auf eine Therapieunzufriedenheit oder den Informationswunsch von Menschen ohne Kontakt zum Gesundheitssystem hindeuten [30]. Bei einer Analyse zu *Pruritus* wurde ersichtlich, dass Berlin und München verglichen mit anderen Großstädten ein besonders hohes Suchvolumen zu analem Juckreiz aufwiesen [29].

» Der Zugang zur medizinischen Versorgung scheint das Suchverhalten der Bevölkerung zu beeinflussen

Zusammenfassend lässt sich sagen, dass die Ergebnisse aus den regionalen Analysen dazu dienen können, die Interessen und insbesondere die medizinischen Bedürfnisse für die unterschiedlichen Regionen zu studieren und zu identifizieren. Demnach könnten so die national und lokal vorhandenen Ressourcen für die medizinische Versorgung und die Bereitstellung von Informationen dort eingesetzt werden, wo sie besonders benötigt werden.



## Fallstricke und Lösungsansätze

Naturgemäß werden im Rahmen der Internetsuchmaschinenanalyse nur Menschen berücksichtigt, die Zugang zum Internet haben und die entsprechende Suchmaschine nutzen. Auch wenn 95 % der internetnutzenden Bevölkerung in Deutschland (90 %, also knapp 63 Mio. Bürger\*innen) Google verwenden [12], werden durch die ausschließliche Berücksichtigung der vom Anbieter erhobenen Daten Informationen über einen Teil der Bevölkerung nicht mit einbezogen. Die gelieferten Daten sind zudem mit Einschränkungen verbunden, da keine demografischen Daten über die Nutzer vorhanden sind. Schließlich ist der Anteil der Onliner unter den 14- bis 59-Jährigen höher als unter den über 59-Jährigen. Auch wenn in den letzten Jahren eine Zunahme bei den über 60-Jährigen verzeichnet wurde (2019: 85 % der 60- bis 69-Jährigen) [4] und bereits gezeigt wurde, dass viele Menschen auch für Familienmitglieder online nach Gesundheitsinformationen suchen [7], könnten Krankheitslasten unter- bzw. überschätzt und altersspezifische Bedürfnisse und Interessenschwerpunkte übersehen werden. Gleichmaßen könnten geschlechtsspezifische Interessen unbeachtet bleiben. Zwar nutzen in etwa gleich viele Frauen und Männer das Internet [4], dennoch kann das Suchverhalten durch geschlechtsspezifische Unterschiede bezüglich Gesundheit und Krankheit (z.B. Unterschiede im Gesundheitsverhalten und in der Wahrnehmung, Bewertung und Kommunikation von Symptomen und gesundheitlichen Beeinträchtigungen) [21] beeinflusst werden. Angesichts dessen wurden in einigen Arbeiten das Durchschnittsalter sowie der Frauenanteil bei der Interpretation von regionalen Unterschieden im Suchverhalten betrachtet, um die Heterogenität zu berücksichtigen.

Da in den aufgeführten Suchmaschinenstudien bis zum jetzigen Zeitpunkt nur Suchanfragen für Deutschland in deutscher Sprache erfasst wurden, können nur Aussagen für deutschsprachige Nutzer getroffen werden. Bisherige Arbeiten, die den Anteil der Personen mit Migrationshintergrund berücksichtigten, konnten allerdings keinen Zusammenhang zwischen einem niedrigeren Suchvolumen und einem

höheren Anteil an Personen mit Migrationshintergrund erkennen [20, 29, 31]. Angelehnt an eine Arbeit von Pereira et al. [19], die Internetsuchanfragen zum Thema *Pruritus* in Europa untersuchten, könnten zukünftig allerdings die Keywords in unterschiedlichen Landessprachen eingesetzt werden. Damit könnten potenzielle Bedürfnisse der vulnerablen Gruppen, die aufgrund ihrer Herkunft oder sprachlicher Defizite Schwierigkeiten beim Zugang zur Gesundheitsversorgung in Deutschland haben [3], identifiziert und gezielt angesprochen werden. Zusätzlich haben bisherige Analysen aufgezeigt, dass auch die medizinische Versorgungsstruktur, die regionale Struktur (ländlich oder urban geprägt) und klimatische Bedingungen das Suchverhalten zu beeinflussen scheinen und bei der Analyse und der Ergebnisinterpretation beachtet werden müssen.

Die Suchanfragen müssen nicht unbedingt von Betroffenen stammen, sondern können auch vom medizinischen Personal, Medizinstudierenden oder Angehörigen, die sich ebenfalls für das Thema interessieren, sein. In Städten mit einem hohen Anteil an Ärzt\*innen oder Medizinstudierenden könnte dies also eine weitere Erklärung für ein höheres Suchvolumen oder spezifischere Suchanfragen sein. Außerdem gibt es keine Informationen darüber, wie häufig dieselben Personen zu einem Thema gesucht haben. Bei der Identifikation von Interessenschwerpunkten und Bedürfnissen ist dies nebensächlich. Bei Explorationen bezüglich der Krankheitslast kann es jedoch zu einer Überschätzung kommen. Mithilfe von für bestimmte Erkrankungen gegenwärtig vorhandenen Surveillance- und Registerdaten können die Internetdaten dahingehend geprüft werden.

Des Weiteren muss berücksichtigt werden, dass das Suchvolumen von *Google Ads Keyword Planner* nur auf Schätzungen eines Google-Algorithmus ohne weitere Informationen darüber, wie das Suchvolumen berechnet wird, basiert. Schließlich muss bei der Interpretation der Daten beachtet werden, dass die Google-Suchmaschine bei der Eingabe eines Suchbegriffes automatisch Suchvorschläge aufführt, die das Suchverhalten der Nutzer\*innen beeinflussen können. Häufiger gesuchte Schlüsselwörter werden mit

größerer Wahrscheinlichkeit vorgeschlagen als seltener gesuchte.

## Das große Ganze

Die Analyse von Suchmaschinendaten hat gegenüber herkömmlichen Verfahren den Vorteil, dass die Daten in nahezu Echtzeit, flächendeckend und ohne großen finanziellen Aufwand auf unterschiedlichen regionalen Ebenen zur Verfügung stehen – was durch die herkömmlichen Datenquellen nur eingeschränkt möglich ist. Die Internetdaten ermöglichen die Identifikation von Interessen und Bedürfnissen einer großen Bevölkerungsgruppe im zeitlichen Verlauf sowie die Überwachung und Bewertung von Krankheitstrends auf unterschiedlichen geografischen Ebenen, die im klinischen Alltag nicht in gleicher Weise erkennbar sind und so z.B. in Notfallsituationen, in denen schnelle datenbasierende Entscheidungen notwendig sind, eingesetzt werden können. Der richtige Gebrauch von Analysen der Suchanfragen kann demnach die optimale Allokation von Ressourcen im Gesundheitssystem maßgeblich fördern und wesentlich dazu beitragen, die Gesundheitsversorgung von Gesundheitsförderung bis Nachsorge zu verbessern, besonders da, wo herkömmliche Daten begrenzt oder nicht vorhanden sind.

## » Der richtige Gebrauch von Analysen der Suchanfragen kann die Gesundheitsversorgung verbessern

Naturgemäß ist die Betrachtung einer einzigen Datenquelle mit Fehlern behaftet. Deshalb ist trotz des enormen Potenzials, das hinter dieser Methode steckt, eine isolierte Betrachtung der Internetsuchmaschinendaten als Gesundheitsdaten zum jetzigen Zeitpunkt nicht zielführend. Für die richtige Interpretation der Daten als „Gesundheitsdaten“ bedarf es daher weiterer Forschungsarbeit, die beispielsweise untersucht, wie sich das Suchverhalten durch einen Besuch beim ärztlichen Fachpersonal verändert.

Insgesamt wird klar, dass durch die Digitalisierung ein schier unendlicher Datenschatz mit einem weitreichenden Potenzial vorhanden ist. Denn nicht nur Suchmaschinendaten bieten die Möglichkeit, Informa-

tionen über Gesundheit und Krankheit in nahezu Echtzeit zu gewinnen. Crowdsourced Data, Informationen, die über partizipative Online-Plattformen, soziale Medien, Apps und Wearables sowie innovative Sensoren [17, 27] gesammelt werden, spielen mittlerweile eine essenzielle Rolle und werden zur Analyse v. a. von Infektionskrankheiten [10, 26] verwendet und bieten völlig neue Möglichkeiten, die Dynamik von Erkrankung bzw. Daten zur Gesundheit in Form von ständig aktuellen digitalen Karten zu untersuchen. Große Möglichkeiten birgt aber auch die Digitalisierung im Gesundheitswesen selbst. Überdies wurde in den letzten Jahren in der Medizin auch zunehmend die Nützlichkeit von Netzwerkanalysen [2] zur Visualisierung und Analyse von komplexen Daten jeglicher Art erkannt. Ziel ist es, diese Daten unter Einhaltung der rechtlichen Grundlagen zusammenzubringen und zu analysieren. Vielversprechend ist hier eine neuartige Vorgehensweise – die „Schwarm-Intelligenz“ –, die eine spezielle Form des Informationsaustausches über verschiedene Knoten hinweg mit Methoden aus dem Werkzeugkasten des „maschinellen Lernens“, einem Teilbereich der künstlichen Intelligenz, kombiniert [36]. So kann der Fortschritt im Gesundheitswesen maßgeblich beschleunigt und die flächendeckende, personenzentrierte Gesundheitsversorgung und der Erhalt der Gesundheit können nachhaltig gestärkt werden.

#### Fazit für die Praxis

- Internetsuchmaschinenanalysen können ein räumlich-zeitliches Abbild hinsichtlich des Interesses und der medizinischen Bedürfnisse der internetnutzenden Bevölkerung liefern.
- Unter der Berücksichtigung potenzieller Fallstricke könnten Suchmaschinen-daten perspektivisch zur Allokation von Ressourcen und der Förderung der Gesundheitsversorgung der Bevölkerung herangezogen werden.
- Um Suchmaschinen-daten als „Gesundheitsdaten“ optimal einsetzen zu können, muss das Suchverhalten der Bevölkerung weiter untersucht werden.
- Ziel sollte es sein, alle verfügbaren Daten zu Erkrankungen zusammenzuführen und im Sinne einer personenzentrierten Medizin zu nutzen.

## Web search data as health data? Geographic differences, temporal trends, and topics of interest from internet search engine analyses in Germany

**Background:** Limited timely, cost-effective, and regional availability of healthcare data is considered one of the limiting factors for timely analyses in the field of health services research and thus for healthcare delivery to the general population. Thus, Internet search engine analyses are coming into focus in this context.

**Objectives:** What contribution can data on the volume of Internet searches for various diseases make to the health care of the German population? Where is the potential and where are the limitations?

**Methods:** A review of the current literature and a selective overview of Internet search engine analyses of dermatological, allergological, and infectious diseases published in the last 4 years in Germany were performed.

**Results:** By analyzing web search data, temporal developments such as seasonality, months with the highest interest, and geographical differences can be mapped at the national, state (*Bundesländer*), and city level and major points of interest (e.g., skin cancer of the eye or anal itching) regarding a certain health topic can be identified. In addition, some studies were able to show correlations between search volume and external factors (e.g., temperature, care structure) and to registered cases (e.g., skin cancer, Lyme disease).

**Conclusion:** Web search data, given the limitations noted, are able to directly provide spatiotemporal mapping regarding the needs of the Internet-using population. They are particularly useful in situations where traditional health data are limited or unavailable.

#### Keywords

Geography · Medical needs · Google · Dermatology · Infectiology

#### Korrespondenzadresse

##### PD Dr. Dr. med. A. Zink, MPH

Technische Universität München, Fakultät für Medizin, Klinik und Poliklinik für Dermatologie und Allergologie am Biederstein  
Biedersteiner Str. 29, 80802 München, Deutschland  
alexander.zink@tum.de

**Förderung.** Keine Förderung

#### Einhaltung ethischer Richtlinien

**Interessenkonflikt.** S. Ziehfreund gibt an, dass kein Interessenkonflikt besteht. L. Tizek erhielt Referentenhonorar und/oder erhielt Förderungen von Beiersdorf Dermo Medical, Janssen Cilag und Novartis. A. Zink war als Berater tätig und/oder erhielt Referentenhonorar und/oder erhielt Förderungen und/oder nahm an klinischen Studien von folgenden Firmen teil: AbbVie, Ammiral, Amgen, Beiersdorf Dermo Medical, Bencard Allergie, BMS, Celgene, Eli Lilly, GSK, Janssen Cilag, Leo Pharma, Miltenyi Biotec, Novartis, Pfizer, Sanofi-Aventis, Takeda Pharma.

Für diesen Beitrag wurden von den Autoren keine Studien an Menschen oder Tieren durchgeführt. Für die aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

#### Literatur

1. Badura-Lotter G (2014) Sexually transmitted diseases: the impact of stigma and taboo on current medical care. *Urologe A* 53:548–556
2. Bauerdorf F, Kaczmarczyk R, Zink A et al (2020) Trends and perspectives for dermatological research in europe: an abstract title analysis of ESDR and IID congresses 2010–2019. *J Invest Dermatol* 140:S197–S200
3. Beaten R, Spasova S, Vanhercke B et al (2018) Inequalities in access to healthcare. A study of national policies. European Union, Brüssel
4. Beisch N, Koch W, Schäfer C (2019) Aktuelle Aspekte der Internetnutzung in Deutschland. ARD/ZDF-Onlinestudie 2019: Mediale Internetnutzung und Video-on-Demand gewinnen weiter an Bedeutung. *Media Perspekt* 9:374–388
5. Bour C, Ahne A, Schmitz S et al (2021) The use of social media for health research purposes: scoping review. *J Med Internet Res* 23:e25736
6. Bujnowska-Fedak MM, Waligora J, Mastalerz-Migas A (2019) The Internet as a source of health information and services. *Adv Exp Med Biol* 1211:1–16
7. Directorate-General for Communication Networks (2014) European citizens' digital health literacy. European Union, Brüssel
8. Gothe H, Swart E, Ihle P (2020) Datennutzung im Gesundheitswesen aus Sicht der Versorgungsforschung. *GGW Wissenschaftsforum Gesundh Ges* 3:7–13
9. Haschke C, Westrick M, Schweck U (2018) Gesundheitsinfos. Wer sucht, der findet –

- Patienten mit Dr. Google zufrieden. Daten Analysen Perspekt 2:1–8
10. Hemler EC, Korte ML, Lankoande B et al (2021) Design and field methods of the ARISE network COVID-19 rapid monitoring survey. *Am J Trop Med Hyg.* <https://doi.org/10.4269/ajtmh.20-1618>
  11. Hilker C, Tizek L, Rütth M et al (2021) Sarcoidose in Germany—Leveraging internet search data to assess prevalence, interest, and unmet needs. *Sci Rep.* <https://doi.org/10.1038/s41598-021-00131-x>
  12. <https://De.Statista.Com/Statistik/Daten/Studie/36146/Umfrage/Anzahl-Der-Internetnutzer-in-Deutschland-Seit-1997/>. Zugriffen: 28. Sept. 2021
  13. Ifsg (2020) „Infektionsschutzgesetz vom 20. Juli 2000 (BGBl. I S. 1045), das zuletzt durch Artikel 12 des Gesetzes vom 10. September 2021 (BGBl. I S. 4147) geändert worden ist“ Änderung durch Art. 12 G v. 10. Sept. 2021 I 4147 (Nr. 63) textlich nachgewiesen, dokumentarisch noch nicht abschließend bearbeitet.
  14. Jing W, Tizek L, Rütth M et al (2021) The national burden of scabies in Germany: A population-based approach using internet search engine data. *PLoS ONE* (under review)
  15. Kammrath Betancor P, Tizek L, Zink A et al (2021) Estimating the incidence of conjunctivitis by comparing the frequency of Google search terms with clinical data: retrospective study. *JMIR Public Health Surveill* 7:e22645
  16. Mick A, Tizek L, Schielein M et al (2021) Can crowdsourced data help to optimize atopic dermatitis treatment? Comparing web search data and environmental data in Germany. *J Eur Acad Dermatol Venereol* (under review)
  17. Miotto R, Danieleto M, Scelza JR et al (2018) Reflecting health: smart mirrors for personalized medicine. *NPJ Digit Med* 1:62
  18. Oteros J, Weber A, Kutzora S et al (2020) An operational robotic pollen monitoring network based on automatic image recognition. *Environ Res* 191:110031
  19. Pereira MP, Ziehfrennd S, Rueth M et al (2021) Google search trends for itch in Europe: a retrospective longitudinal study. *J Eur Acad Dermatol Venereol* 35:1362–1370
  20. Pilz AC, Tizek L, Rütth M et al (2021) Interest in sexually transmitted infections: analysis of web search data terms in eleven large German cities from 2015 to 2019. *Int J Environ Res Public Health* 18:1661–7827
  21. Koch-Institut R (2020) Gesundheitliche Lage der Frauen in Deutschland. Gesundheitsberichterstattung des Bundes. RKI, Berlin (Gemeinsam getragen von RKI und Destatis)
  22. Scheerer C, Rütth M, Tizek L et al (2020) Googling for ticks and Borreliosis in Germany: nationwide Google search analysis from 2015 to 2018. *J Med Internet Res* 22:e18581
  23. Schielein MC, Tizek L, Seifert F et al (2019) Health care of chronic inflammatory skin diseases : do affected individuals seek dermatological care? *Hautarzt* 70:875–882
  24. Schuster B, Tizek L, Schielein MC et al (2021) Retracing the COVID-19 pandemic in Germany from a public perspective using Google search queries related to „coronavirus“. *Gesundheitswesen* 83:e9–e14
  25. Seidl S, Schuster B, Rütth M et al (2018) What do Germans want to know about skin cancer? A nationwide Google search analysis from 2013 to 2017. *J Med Internet Res* 20:e10327
  26. Skinner A, Toumpakari Z, Stone C et al (2020) Future directions for integrative objective assessment of eating using wearable sensing technology. *Front Nutr* 7:80
  27. Someya T, Amagai M (2019) Toward a new generation of smart skins. *Nat Biotechnol* 37:382–388
  28. Swart E, Gothe H, Geyer S et al (2015) Good practice of secondary data analysis (GPS): guidelines and recommendations. *Gesundheitswesen* 77:120–126
  29. Tizek L, Schielein M, Rütth M et al (2019) Influence of climate on Google Internet searches for Pruritus across 16 German cities: retrospective analysis. *J Med Internet Res* 21:e13739
  30. Tizek L, Schielein M, Tizek L et al (2021) Atopische Dermatitis – Identifikation von Bedürfnissen der deutschen Bevölkerung mittels Internetsuchanfragen. *Hautarzt* (under review)
  31. Tizek L, Schielein MC, Rütth M et al (2019) Interest in skin cancer in urban populations: a retrospective analysis of Google search terms in nine large German cities. *Acta Derm Venereol* 99:797–804
  32. Verordnung (EU) (2016) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung). EU, Amtsblatt der Europäischen Union
  33. Wagner N, Zink A, Hell K et al (2021) Patients with chronic urticaria remain largely undertreated: results from the DERMLINE Online survey. *Dermatol Ther* 11:1027–1039
  34. Wallnhöfer F, Erbas M, Tizek L et al (2021) Identifying relevant disease characteristics regarding psoriasis in Germany: a retrospective web search data analysis. *J Med Internet Res* (under review)
  35. Wang R, Helf C, Tizek L et al (2020) The impact and consequences of SARS-CoV-2 pandemic on a single university dermatology outpatient clinic in Germany. *Int J Environ Res Public Health* 17:1660–4601
  36. Warnat-Herresthal S, Schultze H, Shastry KL et al (2021) Swarm Learning for decentralized and confidential clinical machine learning. *Nature* 594:265–270
  37. Zink A (2017) Non-melanoma skin cancer : pathogenesis, prevalence and prevention. *Hautarzt* 68:919–928
  38. Zink A, Rütth M, Schuster B et al (2019) Pruritus in Germany—a Google search engine analysis. *Hautarzt* 70:21–28
  39. Zink A, Schuster B, Rütth M et al (2019) Medical needs and major complaints related to pruritus in Germany: a 4-year retrospective analysis using Google AdWords Keyword Planner. *J Eur Acad Dermatol Venereol* 33:151–156

## MED UPDATE SEMINARE

# 2022

### Derma Update 2022

16. Dermatologie-Update-Seminar

11.–12. November 2022

Berlin

25.–26. November 2022

Mainz

### Wiss. Leitung:

Prof. Dr. Thomas Schwarz, Kiel

Prof. Dr. Thomas Werfel, Hannover

Prof. Dr. Carola Berking, Erlangen

[www.derma-update.com](http://www.derma-update.com)

### Auskunft für alle Update-Seminare:

med update GmbH

[www.med-update.com](http://www.med-update.com)

Tel.: 0611 - 736580

[info@med-update.com](mailto:info@med-update.com)

medupdate