

ასაკის როლი პაციენტთა სამედიცინო ისტორიაში

აბსტრაქტი

ფეიფერში მოცემულია მონაცემთა ანალიზი, რომელიც გვჩვენებს ჯანმრთელობის მდგომარეობის, ასაკისა და სქესის ურთიერთკავშირების გამოსავლენად პაციენტთა დიდი რაოდენობის საშუალებით. ჩვენ ვსწავლობთ ბრაზილიაში 1,7 მილიონი პაციენტის მასიურ ჰეტეროგენულ ნიმუშს, რომელიც შეიცავს 47 მილიონ ჯანმრთელობის ჩანაწერს. სამედიცინო დაწესებულებებში ვიზიტების დეტალური სამედიცინო ჩანაწერები შეგროვებულია 17 თვის განმავლობაში. შედეგები გვკარნახობენ, რომ სამედიცინო პირობები შესაძლებელია დაჯგუფდეს კლასტერებად, რომლებიც იზიარებენ საერთო სიმკვრივეს (განაწილებას) პაციენტთა ასაკში. თითოეული კლასტერისთვის ჩვენ შემდგომში წარმოგიდგენთ ICD-10 დაავადების განაწილებას მათში.

შესავალი

დაავადებათა ჯგუფების ერთობლივი შესწავლა, ან დაავადების თანმხლები დაავადებები, ტრადიციულად ორიენტირებულია დაავადებათა მცირე ჯგუფების კვლევებზე ჰიპოთეზის ტესტირების ტექნიკის გამოყენებით. ცალკეული თანმხლები დაავადებების განმეორებითი არსებობა მნიშვნელოვანია დიაგნოზის დასადგენად. ელექტრონული სამედიცინო ჩანაწერების მონაცემთა ბაზები შეიცავს ფენოტიპურ ინფორმაციას ადამიანების - კერძოდ, პაციენტების კლინიკური ისტორიის შესახებ. ჯანმრთელობის ჩანაწერების ანალიზის ახალი მეთოდი დაფუძნებულია ადამიანის ფენოტიპური დაავადებების ქსელზე, სადაც წევრობები წარმოადგენენ დაავადებებს და წიბოები მიუთითებენ კომორბიდურ ურთიერთობებზე. უფრო ახალი კვლევები კი აანალიზებენ მონაცემთა ბაზებს ჯანმრთელობის ელექტრონული ჩანაწერების შესახებ, რათა აღმოაჩინონ სისტემატური ასოციაციები ცნობილ დაავადებებში. ამ კონტექსტში, ინფორმაციული მეცნიერებების რამდენიმე მეთოდი შეიძლება გამოყენებულ იქნას პაციენტების ელექტრონულ ჩანაწერებში არსებული შაბლონების გამოსავლენად. ამ კვლევების მთავარი ინტერესი არის კორელაციების აღმოჩენა დაავადებებში, რაც ხელს შუწყობს მათ პრევენციას.

დღემდე, ლიტერატურაში შესწავლილი პაციენტების ელექტრონული ჩანაწერების ნიმუშების უმეტესობას აქვს გამოყენება პაციენტების ზოგადი პოპულაციის ვიწრო ნაკრებისთვის. მაგალითად, ჰიდალგომ (Hidalgo)¹ და სხვებმა გამოიყენეს 3 წლის ინფორმაცია 65 წელზე უფროსი ასაკის პაციენტების სამედიცინო მოვლის ისტორიის და ეს იყო მიკერძოებული ინფორმაცია ხანდაზმულთა

მოსახლეობის მიმართ. მოგვიანებით, როქმა (Roque)² და სხვებმა გამოიყენეს პაციენტთა სტრატეგიკაციისა და დაავადებების თანამოხვედრის სტატისტიკა სანქტ ჰანსის ჰოსპიტალიდან, რომელიც არის დანიის უდიდესი ფსიქიატრიული დაწესებულება. მათი შედეგები ფოკუსირებული იყო ფენოტიპებზე, რომლებიც დაკავშირებული იყო ფსიქიკური და ქცევითი დარღვევებთან. თანდათან უფრო ხელმისაწვდომი გახდა მონაცემთა ნაკრები პოპულაციის უფრო სრულყოფილი ნიმუშით. ელექტრონულმა ჩანაწერების დახმარებით ათწლეულების მერე შესაძლებელი გახდა პირველად ავადმყოფების ადრეულ ასაკში გამოვლენა. ჩმიელმა³ და სხვებმა გააანალიზეს ავსტრიაში მთელი მოსახლეობის ორწლიანი სამედიცინო ისტორიები და გაზომეს, თუ როგორ ცვლის კომორბიდობის ქსელი თავის სტრუქტურას ასაკის მატებასთან ერთად პაციენტებში, ეს ინფორმაცია გამოიყენეს დიფუზიური მოდელის შესაქმნელად, რომელიც ხსნის თუ რატომაა დიდ განსხვავება პოპულაციაში დაავადების ვარიაციებს შორის. ამ შემთხვევაში, გაანალიზებული პაციენტების ასაკობრივი ინტერვალების წინასწარ განსაზღვრისას აიგო კომორბიდობის ქსელები. ამ ფეიფერში ჩვენ წარმოადგინეთ კლასტერიზაციის მეთოდს ასაკობრივ სიმკვრივის მსგავსებით ფენოტიპური ჩანაწერები. ჩვენ ვპოულობთ სამედიცინო მდგომარეობების ჯგუფებს, რომლებიც წარმოიქმნება მონაცემებიდან. ეს ჯგუფები თავის მხრივ ასოცირდება ICD-10 ჯგუფებთან.

ექიმების სიბრძნეზე, როდესაც საქმე ეხება დაავადებებისადმი მგრძნობელობის შეფასებას, გავლენა მოახდინა მრავალწლიანმა პრაქტიკამ და ყოველდღიურად ბევრ შემთხვევაზე დაკვირვებამ. ექიმების ცოდნა სხვადასხვა ასაკისა და სქესის დაავადებებისადმი მიდრეკილების შესახებ ეხმარება მათ ჩაატარონ შემოსული პაციენტების დიაგნოსტიკა. პაციენტებისთვის მსგავსმა სიმპტომებმა შეიძლება გამოიწვიოს განსხვავებული დიაგნოზი პაციენტის ასაკისა და სქესის მიხედვით. მაგალითად, 70 წლის პაციენტი ბევრად უფრო ხშირად განიცდის გულის შეტევას, ვიდრე 10 წლის ასაკში, მაშინაც კი, თუ ორივე პაციენტი განიცდის გულმკერდის ტკივილის სიმპტომს. ჩვენ აქ ვაჩვენებთ, რომ ამ შედეგების მიღება შესაძლებელია ჩვენი მონაცემებიდანაც. გარდა იმ სიმპტომებისა, რაც პაციენტს აქვს, ასაკი და სქესი შეიძლება დაეხმაროს დიაგნოსტიკურ პროცესს. ჩვენ წარმოგიდგენთ ფრეიმვორკს, რომელიც ავტომატურად ავლენს ურთიერთობას ჯანმრთელობას მდგომარეობასა და პაციენტის ასაკ/სქესს შორის. ამ მიზნით, ჩვენ ვაჯგუფებთ ჯანმრთელობის მდგომარეობას მოსახლეობის ასაკობრივ სიმჭიდროვესთან მიმართებაში.

მასალები და მეთოდები

1.7 მილიონი პაციენტიდან თითოეულისთვის არის ჟურნალი ექიმთან ყოველი ვიზიტისთვის 17 თვის განმავლობაში 2013 წლის მარტიდან 2014 წლის ივლისამდე. მონაცემები შეესაბამება სამედიცინო ჩანაწერებს ერთ-ერთი უდიდესი ჯანდაცვის სადაზღვევო კომპანიისას ბრაზილიაში. თითოეულ ჯანმრთელობის ჩანაწერი მონაცემთა ბაზაში აქვს რამდენიმე ატრიბუტი, რომელიც ეხება ვიზიტის მონაცემებს, ის სინთეზირებულია ICD-10 კოდების საშუალებით, რომლებიც დეტალურად აღწერს ვიზიტის მდგომარეობას და მიზანს. ICD-10 კოდების აქვთ 23000 სხვადასხვა იდენტიფიკატორი, თითოეული წარმოადგენს პაციენტის ჯანმრთელობის მდგომარეობას. გარდა ამისა, მონაცემებს აქვთ თითოეული პაციენტის ასაკი და სქესი. ვიზიტების საერთო რაოდენობა 6,6 მილიონია, ჯამში 47 მილიონი სამედიცინო ჩანაწერი.

დანართი #1-ზე მოცემულ A ნახაზზე ჩვენ ვაჩვენებთ პაციენტთა მთელი პოპულაციის ასაკობრივ განაწილებას მონაცემებში. ვხედავთ, რომ ასაკობრივი განაწილება პიკს აღწევს 30-40 წელს, რაც არის მედიანური ასაკი, თან ეს მედიანური ასაკი შეესაბამება ბრაზილიის მთელ მოსახლეობას.

B გრაფიკი კი გვიჩვენებს მონაცემებში ICD-10 კოდების სიხშირეების კუმულაციურ განაწილებას. ICD-10 კოდების დაახლოებით 50%-ს ჰქონდა სიხშირე 100-ზე ნაკლების ტოლი პაციენტთა ვიზიტებს შორის, ხოლო დანარჩენი 50% კოდები შეადგენს მონაცემთა ჩანაწერების 90%-ზე მეტს.

ICD-10 კოდირების სქემა სტრუქტურირებულია ხის მეშვეობით და ყველაზე ქვედა ზედა შეიცავს 22 თავს. ICD-10-ის თავებს აქვთ საერთო მახასიათებლები, რომლებიც ეხება იმავე ორგანოს, სისტემას ან ვიზიტის ხასიათს. წვეროების ზომა დანართი #2-ში შეესაბამება ICD-10-ის კონკრეტულ თავის განმეორების სიხშირეს ჩვენს მონაცემებში. წვეროებს შორის წიბოების სისქე კი შეესაბამება პაციენტების ჩანაწერებში ერთდროულად ორივე თავის თანხვედრის სიხშირეს. მაგალითად XXI თავს აქვს ყველაზე მაღალი სიხშირე მონაცემებში, იგი აღწერილია როგორც ჯანმრთელობის მდგომარეობის გავლენის ფაქტორები და კონტაქტები ჯანდაცვის სამსახურებთან, როგორცაა რუტინული გამოკვლევების ჩატარება. X თავი შეიცავს რესპირატორულ სისტემასთან დაკავშირებულ მდგომარეობებს, VII არის თვალის და ადნექსის დაავადებები, XIII არის კუნთოვანი სისტემის და შემაერთებული ქსოვილის დაავადებები და ასე შემდეგ. ყველა თავის აღწერა და დაავადების კოდები მოცემულია ICD10-chapters ცხრილში.

გადამოწმებით ჩანს, რომ თითოეულ ICD-10 კოდს აქვს სიმკვრივის გამორჩეული განაწილება ასაკობრივ განზომილებაში, რომელიც მოიცავს სხვადასხვა

ასაკობრივ ჯგუფს დაბადებიდან მოყოლებული. დანართი #3 გვიჩვენებს ჩუტყვავილას და გლაუკომის ასაკობრივი სიმკვრივის ნიშნების მაგალითს. როგორც მოსალოდნელი იყო, გლაუკომა უფრო გავრცელებულია ხანდაზმულებში და ჩუტყვავილა ბავშვებში. ამ განაწილების ფორმები მიუთითებს იმაზე, რომ არსებობს სხვადასხვა დაავადების გარკვეული ასაკის პაციენტებში გამოვლენის უფრო მაღალი ალბათობა.

ჩვენ გამოვიყენეთ იერარქიული კლასტერიზაციის მეთოდი, რათა დაგვენახა გარკვეული კანონზომიერებები სხვადასხვა დაავადების მქონე პაციენტების ასაკობრივ ჯგუფებში. კლასტერიზაციისთვის მონაცემები გარდაქმენით შემდეგნაირად, თითოეული დაავადებისთვის ავაგეთ 1x100 ვექტორი, სადაც ვექტორის თითოეულ პოზიციაზე განთავსებული რიცხვი აღნიშნავს ალბათობას, ამ პოზიციის შესაბამის ასაკში, ამ კონკრეტული დაავადების ქონისა. თავდაპირველად იერარქიული კლასტერიზაცია თითოეულ ვექტორს განიხილავს როგორც დამოუკიდებელ კლასტერს და შემდგომ იწყებს მათ გაერთიანებას, მანამ სანამ არ ჩამოყალიბდება ერთი კლასტერი. ორ ვექტორს შორის მანძილის დასათვლელად ჩვენ გამოვიყენეთ ევკლიდეს მანძილის დათვლის ფორმულა.

იერარქიული კლასტერიზაციის შემდგომ, საჭირო იყო შეგვეჩია კლასტერების რაოდენობა, ამისთვის გამოვიყენეთ ე.წ. „იდაყვის წესი“.

შედეგები

ჩვენ შემდგომ ვაანალიზებთ ICD-10 კოდების ასაკობრივ სიმკვრივებს მონაცემებში, რომ მოვახდინოთ ICD-10 კოდების სეგმენტირებისთვის ასაკის მსგავსი სიმკვრივის მქონე პირობებით დაჯგუფება. ჩვენ განვიხილავთ ანალიზს XVIII-XXII კოდების გამოკლებით. გამორიცხული კოდები მოიცავს სიმპტომებს (მაგ. R კოდები), პროცედურულ დეტალებს (როგორიცაა გართულებები ან წამლის გვერდითი ეფექტები) და ასევე პირადი ფაქტორები (ზოგადი გამოკვლევები და სხვა). ჩვენ წარმოვადგენთ ასაკობრივ განაწილებას ვექტორის სახით, რომელიც შედგება 100 ელემენტისგან და თითოეულ ელემენტს აქვს შესაბამისი ალბათობა თუ რა ნაწილს აქვს ეს კოდი თითოეული ასაკისთვის. ეს განისაზღვრება, როგორც ალბათობა $P(\text{age} | \text{code})$. სადაც age არის პაციენტის ასაკი, ხოლო code არის იმ პაციენტის სიმრავლე, რომელთაც შესაბამისი ICD-10 კოდი აქვთ.

ჩვენ დავაკლასტერეთ სიმკვრივეები თითოეული ICD-10 კოდისთვის ასაკის ვექტორული წარმოდგენის საფუძველზე. ვიყენებთ იერარქიულ აგლომერაციულ

კლასტერიზაციას (HAC) კოდების კლასტერებად დაჯგუფებისთვის. კოდების ასაკობრივი განაწილება ჯგუფდება ექვს ძირითად ჯგუფად, როგორც ნაჩვენებია დანართი #4-ზე. კლასტერები A და B აჩვენებს კოდების ორი კლასტერს, რომლებსაც აქვთ უფრო მაღალი სიმკვრივე დაბალ ასაკში. C კლასტერი გვიჩვენებს კოდების ჯგუფს, რომლებსაც აქვთ ასაკობრივი სიმკვრივე კონცენტრირებული 20-დან 40 წლამდე. E კლასტერს 60 წელზე მეტი ასაკის დიაპაზონში, ხოლო F კლასტერს აქვს ასაკობრივი სიმკვრივის კოდების კონცენტრაცია 70-ზე მეტი წლის შემთხვევაში.

დანართი #5-ზე ნაჩვენებია რამდენიმე მაგალითი ისეთი ICD-10 კოდების, რომელთაც მაღალი განაწილება აქვთ ჩვენს მონაცემებში. ნახაზი გვიჩვენებს კლასტერიზაციის დენდროგრამას, რომლის სიღრმეც 6-ის ტოლია. ტოტები რომელთა სიღრმეც 6-ზე მეტია წარმოდგენილია ისეთი დაავადებით, რომელიც მის შვილებში ყველაზე მეტადაა გავრცელებული და თითოეული მათგანი მონიშნულია კლასტერის ნომრით A-დან F-ის ჩათვლით. A კლასტერში, J21 მწვავე ბრონქიოლიტი და H65 ოტიტი დაფიქსირდა მოსახლეობის 0.4 და 1.2 პროცენტში შესაბამისად, ორივეს აქვს კონცენტრაცია ქვედა ასაკის მიმართ, როგორც ეს ადრე იყო ნაჩვენები. B კლასტერს აქვს ზედა სასუნთქი გზების J06 მწვავე ინფექციები პაციენტების მოსახლეობის 8 პროცენტით. გარდა ამისა, მას აქვს A09 დიარეა და J03 მწვავე ტონზილიტი თითოეული დაახლოებით 5.9 პროცენტის ტოლი. C კლასტერში O82 საკეისრო კვეთით მშობიარობას აქვს პაციენტთა პოპულაციის დაახლოებით 0.8 პროცენტი, ეს კლასტერი შეესაბამება განსაზღვრულ ასაკობრივ დიაპაზონს 20-დან 40 წლამდე როგორც დასაწყისში იყო ნახსენები. D კლასტერს აქვს H52 რეფრაქციის დარღვევები და წარმოდგენილია 10.6 პროცენტით, ხოლო J01 მწვავე სინუსიტი პაციენტების მოსახლეობის 6.7 პროცენტით. როგორც მოსალოდნელი იყო, რამდენადაც კლასტერებს უფრო მეტი სიმჭიდროვე აქვთ პოპულაციის ასაკობრივი განაწილების პიკის გარშემო, პაციენტების რაოდენობა თითოეულ კოდზე კლასტერებში უფრო მაღალი ხდება. E კლასტერს ასაკობრივი სიმკვრივით უმეტესწილად ხანდაზმულებში აქვს M54 ზურგის ტკივილი 10.8 პროცენტით და M25 სახსრების სხვა დარღვევები 4.7 პროცენტით. კლასტერი F ასაკობრივი სიმკვრივით მიეკუთვნება ყველაზე ხანდაზმულ ჯგუფს და აქვს I10 ესენციური ჰიპერტენზია 10.4 პროცენტში და N39 სამარდე გზების სხვა დარღვევები დაახლოებით 3.5 პროცენტში. პნევმონია მესამეა დაახლოებით 1.8 პროცენტით წარმოდგენილი.

დასკვნა

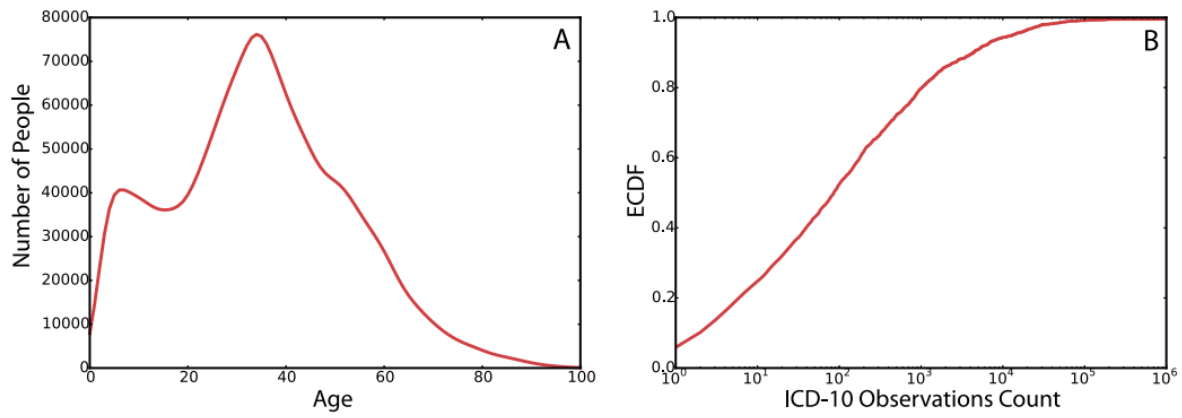
ჩვენი თემის მიზანი იყო, გაგვეჩვენა პაციენტის ასაკის დამოკიდებულება სხვადასხვა დაავადებებთან. როგორც გაირკვა, სიმპტომების გარდა დიაგნოზის დასმისას პაციენტის ასაკიც განსაკუთრებულ როლს თამაშობს. იერარქიული კლასტერიზაციის მეშვეობით, ჩვენ გამოვყეთ ექვსი ძირითადი ჯგუფი, რომლებიც მოიცავენ მსგავსი ტიპის დაავადებებს. რადგან კლასტერიზაციის შედეგად აღმოვაჩინეთ, რომ ასაკი არის დამოკიდებულებაში პაციენტის დაავადებებთან, მიგვაჩნია რომ შესაძლებელია პაციენტის სიმპტომების, ასაკის და სქესის გათვალისწინებით მოხდეს დაახლოებითი პირველადი დიაგნოზის განსაზღვრა ავტომატიზაციით.

ბიბლიოგრაფია

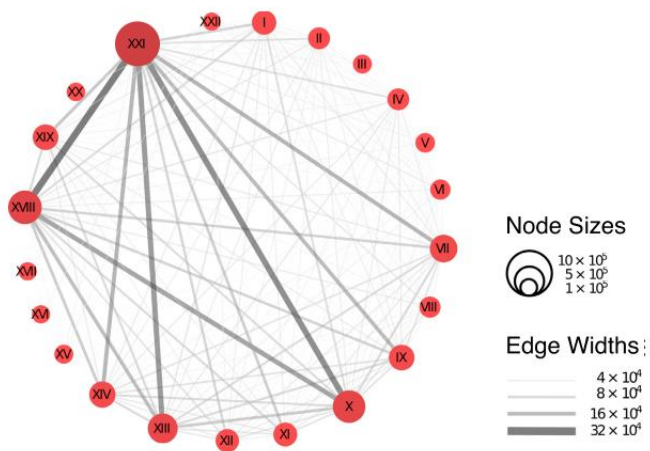
1. Hidalgo CA, BlummN,Baraba'si AL, Christakis NA. A dynamicnetworkapproach for the study of humanphenotypes. PLoSComputBiol.2009;5(4): e1000353. <https://doi.org/10.1371/journal.pcbi.1000353> PMID:19360091
2. RoqueFS,JensenPB,SchmockH,DalgaardM,AndreattaM,HansenT,etal.Usingelectronic patient records to discover disease correlations and stratify patient cohorts. PLoS Comput Biol. 2011; 7(8): e1002141. <https://doi.org/10.1371/journal.pcbi.1002141> PMID: 21901084
3. ChmielA,KlimekP,ThurnerS.Spreadingofdiseasesthroughcomorbidity networksacrosslife and gender. NewJournalof Physics. 2014; 16(11):115013. <https://doi.org/10.1088/1367-2630/16/11/ 115013>

დანართი

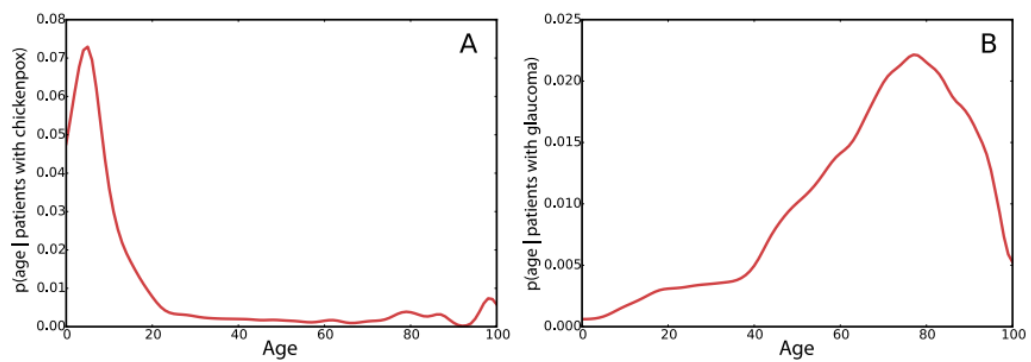
დანართი #1



დანართი #2



დანართი #3



დანართი #4

