

Haigusregistrite andmekvaliteedi uurimise meetodid

Katrin Lang – Tartu Ülikooli tervishoiu instituut, Londoni Hügieeni ja Troopilise Meditsiini Kool

andmekvaliteet, kvaliteediuring, register

Rahvastikul põhinevatesse haigusregistritesse kogutakse epidemioloogilisi andmeid haiguste esinemissageduse kohta. Täiuslik register peaks sisaldama iga registreerimisele kuuluva juhu kohta ühe võimalikult täpse ja täiusliku kirje. Tegelikuses ilmneb andmekvaliteedis puudujääke. On ülimalt tähtis teada, millised on puudujäägid registriandmete täielikkuses ja täpsuses, kuna neid andmeid kasutatakse laialdaselt kliinilistes ning epidemioloogilistes uurimustes. Artiklis on antud ülevaade meetoditest, mille abil saab hinnata haigusregistrite andmekvaliteeti. Andmekvaliteedi tavauuringud peaksid olema iga registri igapäevatöö osa. Lisaks sellele oleks aga vaja aeg-ajalt korraldada eriuuringuid, et saada väärtuslikku lisateavet andmekvaliteedi kohta.

Rahvastikupõhised haigusregistrid koguvad andmeid haiguste esinemissageduse kohta. Registreeritud andmete väärtus sõltub suuresti nende kvaliteedist ja seetõttu on väga tähtis, et andmed oleksid täielikud ja täpsed (1–2). Registrite andmeid kasutatakse laialdaselt epidemioloogilistes ja kliinilistes uuringutes, mille tulemuste valiidsus sõltub suurel määral andmekvaliteedist (3–7). Andmete kvaliteedi saavutamiseks peab register esiteks tagama kõigi haigusjuhtude registreerimise, mis tekivad hõlmatavas rahvastikus ja mis kuuluvad registreerimisele (8). Teiseks peavad registri käsutuses olema täpsed andmed nn riskirahvastiku ehk hõlmatava rahvastiku soo- ja vanuskoostise kohta, et oleks võimalik välja arvutada haigestumuskindrad. Stilleri (9) järgi sisaldab täiuslik register täieliku ja täpse kirje iga juhu kohta. Tegelikuses on seda võimatu saavutada, sest vead võivad tekkida nii juhu kindlakstegemisel kui registreerimisel, samuti võib juht jääda mingil põhjusel registreerimata. Nüüd tekibki küsimus, mil määral esineb haigusregistri andmetes vigu ja puudusi. Iga register peab olema võimeline andma hinnangu andmekvaliteedile (10).

Registri andmekvaliteedi hindamiseks on välja töötatud mitmed meetodid, millest on artiklis antud ülevaade. Varem on eesti keeles vähiregistri kvaliteediküsimusi käsitlenud M. Rahu.

Terminoloogia

Juhtude registreerimise täielikkus (*case completeness of registration or completeness of case ascertainment*) näitab, kui suur osa registri poolt hõlmatavas rahvastikus asetleidnud juhtudest on kantud registrisse (8). **Registreerimise täpsus** (*accuracy of registration*) peegeldab registriandmete vastavust algallikas (s.o tervishoiuasutuste dokumentatsioonis) sisalduvale informatsioonile (11). Seda mõõdetakse tavaliselt kui vastavust kuldstandardile (12).

Peale nimetatud andmekvaliteedi aspektide on veel teisi, näiteks andmete kogumise kiirus (keskmine aeg haiguse diagnoosimisest registreerimiseni), tunnuste kodeerimise püsivus aja jooksul, registriandmetel põhinevate publikatsioonide arvu püsivus (11, 13) ja veel mõned, mida selles põhilisi andmekvaliteedi näitajaid kirjeldavas ülevaateartiklis ei ole käsitletud.

Kvaliteedi kontrolli meetodid on suunatud mingi kvaliteediaspekti uurimisele. Andmekvaliteedi **tavauuringud** kasutavad juba registris või muudes andmebaasides olemasolevaid andmeid. **Eriauuringud** on sellised, kus kogutakse spetsiaalselt selliseid andmeid, mille abil hinnatakse registri andmekvaliteeti.

Juhtude registreerimise täielikkus

Nagu eespool öeldud, on rahvastikupõhise registri eesmärgiks registreerida kõik juhud, mis tekivad tema poolt hõlmatavas rahvastikus ja mis kuuluvad registreerimisele. Selle eesmärgi saavutamine tähendaks 100% registreerimise täielikkust, mida tegelikkuses on peaaegu võimatu saavutada. Kõrvalekalded registreerimise täielikkusest jaotuvad üle- ja alaregistreerimiseks. Üleregistreerimisega on tegemist siis, kui registrisse satuvad kas topeltjuhud või juhud, mis on aset leidnud väljaspool hõlmatavat rahvastikku, või kui on tegemist eksidiagnoosiga (nn väärpositiivse diagnoosiga). Alaregistreerimine tekib siis, kui juht jääb mingil põhjusel registrisse teatamata. Teine põhjus alaregistreerimiseks on see, et haiguste puhul võib olla tegemist aladiagnostikaga ja kui haigusjuhtu ei ole diagnoositud, ei saa seda registreerida.

Juhtude registreerimise täielikkuse tavauuringud on sisestuskontroll, surmatunnistuste meetod ning suremuse ja haigestumuse suhte hindamine. Viimased kaks leiavad põhiliselt rakendust vähiregistrite andmekvaliteedi uurimisel.

Sisestuskontroll (*edit checks*) on andmesisestuse osa. Selle kasutamine aitab vältida topeltjuhtude registreerimist. Topeltjuhtude registreerimise oht tekib siis, kui sama juhu kohta saadetakse registrisse mitu teatist. Sellisel juhul aitab sisestuskontroll vältida sama juhu sattumist registrisse mitu korda. Sisestuskontroll aitab ka vältida nn loogikavigu kuupäevade (nt diagnoosimise kuupäev hilisem kui ravi alguse kuupäev) ja diagnooside (nt eesnäärmevähk naispatsiendil) registreerimisel.

Surmatunnistuse meetod (*death certificate method*) on üks sagedamini kasutatav juhtude registreerimise täielikkuse hindamise meetodeid ning seda käsitatakse igapäevase kvaliteedikontrolli osana. See meetod hindab surmatunnistusjuhtude osakaalu kõigi registreeritud juhtude seas. See tähendab, et varem, enne surmatunnistusel oleva teabe saamist pole vähiregister selle vähijuhu kohta informatsiooni saanud (8). Meetod põhineb sellel, et igal aastal vähiregistri andmebaas lingitakse riigi keskse statistika-

institutsiooni surmatunnistuste failiga. Eesmärgiks on leida need vähijuhud, mille kohta vähiregister informatsioon puudub. Tulemus esitatakse protsendina, mis näitab surmatunnistuse alusel registreeritud vähijuhtude osakaalu kõigist vastaval perioodil registreeritud juhtudest. Siin eristatakse kindlalt ainult surmatunnistuse-juhte ehk DCO (*death certificate only*) juhte.

Suremuse ja haigestumuse suhe (*mortality-incidence ratio*). Sedagi meetodit kasutatakse vähiregistris tavalise andmekvaliteedi kontrolli osana. Arvutatakse suremus- ja haigestumuskordajate suhe (8). Seda näitajat tõlgendatakse nii, et väärtusele üks lähenev näitaja väljendab selle vähiliigiga patsientide väga väikest elulemust ning et suhteliselt suurema elulemuse korral on näitaja ühest oluliselt väiksem. Suremuse ja haigestumuse suhe, mis on ühest suurem, näitab üldjuhul alaregistreerimist; haruharva tuleb arvesse olukord, kus haigestumus on vaadeldaval ajaperioodil kiiresti vähenenud. Meetodi kasutamisel peab arvestama, et diagnoosi puudutav informatsioon ei pruugi surmatunnistusel olla täpne, eriti kui pole tehtud lahangut. Näiteks on täheldatud (14–15), et surmapõhjus on surmatunnistusel antud väiksema täpsusega kui haigusloo lõppdiagnoos.

Juhtude registreerimise täielikkuse uurimise eri-uuringud on ajalooliste andmete meetod, aktiivne päring ning nn püügi-taaspüügi (*capture-recapture*) meetod.

Ajalooliste andmete meetod (*historic data method*) võrdleb registri poolt mingil ajavahemikul registreeritud juhtude arvu selle juhtude arvuga, mis kajastab juhtude teket märkimisväärselt sarnase rahvastiku hulgas sama pika ajavahemiku jooksul (8). Võrdlusandmestikuks võivad olla kas sama registri andmed varasemast ajaperioodist või andmed mõne teise rahvastiku kohta, millel on uuritava registri poolt hõlmatava rahvastikuga sarnane etniline ning vanus- ja sookoostis. Sellel meetodil on hulk puudusi alates sellest, kas ja kuidas oli mõõdetud juhtude registreerimise täielikkus võrdlusrahvastiku andmestikus, ning lõpetades

sellega, et tegelikkuses ei ole võimalik leida kõikide tulemusi mõjutada võivate tunnuste poolest absoluutselt sarnaseid rahvastikke. Meetodi nimetust silmas pidades peaks eelkõige tegema võrdluse sama registri varasemate aastate andmetega, kuid siis on tulemusi moonutavaks asjaoluks haigestumuse muutumine aja jooksul. Meetodi kasutamine annab teavet üksnes võimalike suurte kõrvalekallete kohta registreerimise täielikkuses ning innustab korraldama teisi kvaliteediuuringuid registris.

Aktiivne päring (*active retrieval*) kujutab endast registri andmete kõrvutamist mingi muu andmebaasiga, kus on küllalt hea juhtude registreerimise täielikkus ja kuhu kogutakse andmeid sama rahvastiku seast. Sellisteks andmebaasideks võivad olla mingi teadusuuringu käigus loodud andmebaas, samuti administratiivne või ravikindlustuse andmebaas. Vastavalt haiguse spetsiifikale võib kasutada ka teisi andmebaase, vähiandmete kvaliteedi uurimisel näiteks patoloogia-teatiste andmebaasi. Tulemuste valiidsus onoleb võrdluseks kasutatava andmebaasi täpsusest ja täielikkusest. Andmebaaside võrdlemisel tehakse kindlaks need juhud, mis ei ole kantud antud registrisse, ning nende kohta otsitakse lisainformatsiooni muudest dokumentidest (nt haiguslood). Kui diagnoos leiab sellise aktiivse päringu tulemusena kinnitust, siis kantakse ta registrisse. Juhtude registreerimise täielikkust on teiste andmebaasidega linkimise tulemusena uuritud palju (2, 7, 10, 16–21). Eestis on seda meetodit kasutatud Eesti vähiregistri andmekvaliteedi uurimisel (22).

Viimastel aastatel on juhtude registreerimise täielikkuse uurimisel populaarseks saanud nn **püügi-taaspüügi meetod** (*capture-recapture method*) (8, 23–25). See pärineb loomaökoloogiast, kus hinnatakse looduses elavate loomade arvu. Nimelt püütakse ja märgistatakse üksteisest sõltumatud valimid ning antakse hinnang loomade üldarvule vastavalt sellele, kui palju loomi püüti ja taaspüüti kattuvates valimites. Epidemioloogilises seires kujutavad mittetäieliku juhtude registreerimisega andmebaasid endast kattuvaid valimeid ning seis

juhtude arvu loendamine annab ettekujutuse juhtude üldarvust.

Selle meetodi kasutamiseks peavad juhud olema loomulikult tuvastatavad kas nime või mõne muu tunnuse kaudu (8). Järgnevalt jaotatakse juhud selle järgi, kas nad esinevad või mitte igas vaatluse all olevas andmebaasis. Kui andmebaaside arv on n , siis on kaks n -i võimalikku kombinatsiooni, et klassifitseerida juhte vastavalt sellele, kas nad käsitletavates andmebaasides eksisteerivad või mitte. Näiteks kahe andmebaasi – haiguslugude ja surmatunnistuste puhul on neli võimalust: iga konkreetne juht on olemas nii haiguslugude kui surmatunnistuste andmebaasis, ainult haiguslugude andmebaasis, ainult surmatunnistuste andmebaasis või ei leidu teda kummaski. Mõlemast andmebaasist puuduvate juhtude arv leitakse kolme eelneva kombinatsiooni juhtude arve kasutades. Seda kasutades on võimalik leida kummagi andmebaasi juhtude registreerimise täielikkus.

Selle meetodi puhul kasutatakse kas originaal-andmebaasi (näiteks mingi konkreetse raviasutuse haiguslood) või analüütilist andmebaasi, mille on uurija koostanud ühe või enama andmebaasi alusel (26). Püügi-taaspüügi meetodi kasutamisel tuleb silmas pida, et andmebaasid peavad olema üksteisest sõltumatud (8). See tähendab, et iga üksiku juhu registreerimine ühes vaatluse all olevas andmebaasis ei muuda juhu registreerimise tõenäosust uuringus kasutatavas muus andmebaasis. Kui seda tingimust arvesse ei võeta, siis tulemused kas ala- või ülehindavad puuduvate juhtude arvu.

Püügi-taaspüügi meetodit on laialdaselt kasutatud eri tüüpi registreerimise korral, mis registreerivad vähki (25, 27, 28), diabeeti (29, 30), põletikulist soolehaigust (31), tsöliaakiat (32), insulti (33), fetaalsündroomi (34), epilepsiat (35), tserebraalparalüüsi (36) ja tuberooset skleroosi (37).

Registreerimise täpsus

Registreerimise täpsus peegeldab registriandmete vastavust algallikas (s.o tervishoiuasutuste dokumentatsioonis) sisalduvale informatsioonile.

Ebatäpsused registriandmetes võivad tekkida mitmel viisil: haiguslugudest väljakirjutuste tegemisel, andmete sisestamisel registri andmebaasi või kodeerimisel (38). Registreerimise täpsust võimaldavad hinnata diagnostiliste kriteeriumite meetod ja teatiste uuesti täitmise meetod (39). Registriandmete täpsust mõjutavad mitte ainult vigased, vaid ka puuduvad andmed. Kirjete registreerimise täielikkus viitab registreeritavate tunnuste kaupa puuduvate väärtustega kirjete osatähtsusele kirjete üldarvus.

Diagnostiliste kriteeriumite meetod

(*diagnostic criteria method*) on suunatud registreeritava haiguse diagnoosi täpsuse uurimisele. Ta mõõdab diagnostilistele kriteeriumitele vastavate diagnooside osatähtsust kõigi registris leiduvate diagnooside seas. Vähiregistrites arvutatakse selle meetodi puhul histoloogiliselt (morfoloogiliselt) kinnitatud juhtude protsent.

Teatiste uuesti täitmise meetod

(*reabstraction method*) on registreerimise täpsuse uurimise kaudne meetod, mis põhineb sellel, et moodustatakse juhuvalim uuritavatest juhtudest, mis uuringu käigus algdokumentidest uuesti registreeritakse, kasutades igapäevaselt käibel olevat teatise vormi. Et spetsiaalse uuringu käigus on dokumentide läbivaatamine eeldatavasti põhjalikum kui tavaliselt haiguseteatise täitmise käigus, vaadeldakse uuesti täidetud teatist kui kuldstandardit, millega võrreldakse registris vastavate juhtude kohta olevat informatsiooni tunnuste kaupa. Tulemused avaldatakse kokkulangevuse protsendina. Registriandmete täpsuse uurimiseks on teatiste uuesti täitmise meetodit kasutatud suhteliselt palju (2, 3, 6, 20, 40–42).

Kirjete registreerimise täielikkus (*completeness of data items*) annab ülevaate puuduvatest väärtustest registreeritavate tunnuste kaupa. Alati ei ole võimalik juhtudest teatamisel täita kõikide tunnuste jaoks vastavaid lahtreid teatisel, sest see informatsioon võib puududa registreerimise aluseks olevatel meditsiinidokumentidel. Ka võib dokumentidel olla osa lahtreid täitmata või veelgi halvem, seal puuduvad üldse kohad vastava

informatsiooni registreerimiseks. Teatud registreerimistunnused, näiteks sugu, omavad suuremat tähtsust kui mõned teised, näiteks abieluseis või elukutse. Puuduvate väärtuste hindamisel tuleb arvestada tunnuste tähtsust ning anda neile hinnang lähtuvalt uuritava registri olemusest. Tulemused esitatakse sagedustabelitena, kus näidatakse puuduvate väärtuste protsendid uuritavate tunnuste kaupa. On võimalik vaatluse alla võtta registri tegevuse mingi ajaperiood, et hinnata registreerimise täielikkuse võimalikke muutusi ajas.

Kokkuvõte

Eespool kirjeldatud meetodid annavad ülevaate sellest, kuidas hinnata haigusregistrite andmekvaliteeti. Et vähk on üks tüüpilisemaid ja sagedamaid registreeritavaid haigusi, on paljud meetodid välja kasvanud vähiregistrite andmekvaliteedi hindamisest ning mõned on rakendatavad ainult sel eesmärgil. Mitmeid meetodeid saab kasutada ka muude haigusregistrite või teiste meditsiiniregistrite andmekvaliteedi uurimisel. Andmekvaliteedi tavauuringud peaksid olema registri igapäevatöö osa. Lisaks sellele oleks aga vaja aeg-ajalt korraldada eriuuringuid, et saada väärtuslikku lisateavet andmekvaliteedi kohta.

Tänuavaldus

Autor tänab prof Dave Leoni Londoni Hügieeni ja Troopilise Meditsiini Koolist ning prof Mati Rahu Eksperimentaalse ja Kliinilise Meditsiini Instituudist kasulike kommentaaride ning nõuannete eest käsikirja valmimisel. Autori õpinguid Londoni Hügieeni ja Troopilise Meditsiini Koolis toetas Eesti Tervishoiuprojekt.

Kirjandus

1. Blanchard JF, Dean H, Anderson K, Wajda A, Ludwig S, Depew N. Incidence and prevalence of diabetes in children aged 0–14 years in Manitoba, Canada, 1985–1993. *Diabetes Care* 1997;20:512–5.
2. Schouten LJ, Van der Does-van den Berg, Otter R, Coebergh JW. Accuracy and completeness of the registration of childhood leukaemia in The Netherlands, 1989–1992. *Eur J Cancer* 1997;33(6):891–94.
3. Gulliford MC, Bell J, Bourne HM, Petrukevitch A. The reliability of cancer registry records. *Br J Cancer* 1993;67(4):819–21.
4. Harvei S, Tretli S, Langmark F. Quality of prostate cancer data in the cancer registry of Norway. *Eur J Cancer* 1996;32A(1):104–5.
5. Brewster D. Improving the quality of cancer registration data. *J R Soc Med* 1995;88(5):268–71.
6. Lapham R, Waugh NR. An audit of the quality of cancer registration data. *Br J Cancer* 1992;66(3):552–4.
7. Kjaergaard J, Clemmensen IH, Storm HH. Validity and completeness of registration of surgically treated malignant gynaecological diseases in the Danish National Hospital Registry. *J Epidemiol Biostat* 2001;6(5):387–92.
8. Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL. Comparability and Quality Control in Cancer Registration. Lyon: IARC; 1994.
9. Stiller CA. Reliability of cancer registration data. *Eur J Cancer* 1997;33(6):812–4.
10. Nwene U, Smith A. Assessing completeness of cancer registration in the north-western region of England by a method of independent comparison. *Br J Cancer* 1982;46:635–9.
11. Skeet RG. Quality and quality control. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG. *Cancer Registration: Principles and Methods*. IARC Scientific Publications No 95. Lyon: IARC; 1991. 101–7.
12. Teperi J. Use of registry data in the analysis of medical practices. Helsinki: University of Helsinki, Department of Public Health; 1994.
13. Hilsenbeck SG, Glaefke GS, Feigl P, Lane WW, Golenzer H, Ames C, et al. Quality control for cancer registries. Washington, DC: U.S. Department of Health and Human Services; 1995.
14. Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health* 1981;71:242–50.
15. Cameron HM, McGoogan E. A prospective study of 1152 hospital autopsies: I. Inaccuracies in death certification. *J Pathol* 1981;133:273–83.
16. Brewster DH, Crichton J, Harvey JC, Dawson G, Nairn ER. Benefits and limitations of pathology databases to cancer registries. *J Clin Pathol* 1996;49(11):947–9.
17. Brewster DH, Crichton J, Harvey JC, Dawson G. Completeness of case ascertainment in a Scottish regional cancer registry for the year 1992. *Public Health* 1997;111(5):339–43.
18. Aström M, Bodin L, Tidfelt U. Adjustment of incidence rates after an estimate of completeness and accuracy in registration of acute leukemias in a Swedish population. *Leuk Lymphoma* 2001;41(5–6):559–70.
19. Parkin DM, Wabinga H, Namboozee S. Completeness in an African cancer registry. *Cancer Causes Control* 2001;12:147–52.
20. Dickinson HO, Salotti JA, Birch PJ, Reid MM, Malcolm A, Parker L. How complete and accurate are cancer registrations notified by the National Health Service Central Register for England and Wales? *J Epidemiol Community Health* 2001;55:414–22.
21. Pobereskin LH. The completeness of brain tumour registration in Devon and Cornwall. *Eur J Epidemiol* 2001;17:413–416.
22. Lang K, Mägi M, Tikk M, Aareleid T. Eesti Vähiregistri andmete täielikkuse uuring. *Eesti Arst* 2001;80:320–4.
23. Brenner H. Effects of misdiagnoses on disease monitoring with capture-recapture methods. *J Clin Epidemiol* 1996;49(11):1303–7.
24. Brenner H. Use and limitations of the capture-recapture method in disease monitoring with two dependent sources. *Epidemiology* 1995;6(1):42–8.
25. Brenner H, Stegmaier C, Ziegler H. Estimating completeness of cancer registration in Saarland/Germany with capture-recapture methods. *Eur J Cancer* 1994;30A(11):1659–63.
26. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiol Rev* 1995;17(2):243–64.
27. Hook EB. Capture-recapture estimates of cancer incidence, adjusting for geographic effects: an alternative perspective [letter; comment]. *Am J Public Health* 1996;86(5):746–7.
28. Robles SC, Marrett LD, Clarke EA, Risch HA. An application of capture-recapture methods to the estimation of completeness of cancer registration. *J Clin Epidemiol* 1988;41(5):495–501.
29. Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW et al. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. *BMJ* 1997;315(7107):524–8.
30. Wadsworth E, Shield J, Hunt L, Baum D. Insulin dependent diabetes in children under 5: incidence and ascertainment validation for 1992. *BMJ* 1995;310(6981):700–3.
31. Tragnone A, Corrao G, Miglio F, Caprilli R, Lanfranchi GA. Incidence of inflammatory bowel disease in Italy: a nationwide population-based study. *Int J Epidemiol* 1996;25(5):1044–52.
32. Corrao G, Usai P, Galatola G, Ansaldi N, Meini A, Pelli MA et al. Estimating the incidence of coeliac disease with capture-recapture methods within four geographic areas in Italy. *J Epidemiol Community Health* 1996;50(3):299–305.

33. Taub NA, Lemic-Stojcevic N, Wolfe CD. Capture-recapture methods for precise measurement of the incidence and prevalence of stroke [letter]. *J Neurol Neurosurg Psychiatry* 1996;60(6):696-7.
34. Sorensen HT, Hansen I, Ejlersen E, Sabroe S, Hamburger H. Identification of patients treated with strong analgesics: an assessment of two Danish information systems with respect to epidemiological research. *J Med Syst* 1996;20(1):57-65.
35. Murphy CC, Trevathan E, Yeargin-Allsopp M. Prevalence of epilepsy and epileptic seizures in 10-year-old children: results from the Metropolitan Atlanta Developmental Disabilities Study. *Epilepsia* 1995;36(9):866-72.
36. Topp M, Langhoff-Roos J, Uldall P. Validation of a Cerebral Palsy Registry. *J Clin Epidemiol* 1997;50:1017-23.
37. O'Callaghan FJ, Shiell AW, Osborne JP, Martyn CN. Prevalence of tuberous sclerosis estimated by capture-recapture analysis. *Lancet* 1998;351(9114):1490.
38. Schouten LJ. Cancer registration: data quality and prospects for use. Heythuysen: Datawys / Universitaire Pers Maastricht; 1996.
39. Kuntoro, LaPorte RE, Mazumdar S. Approaches to quality control with an application to a new cancer registry in a developing country. *J Clin Epidemiol* 1994;47(7):779-86.
40. Middleton RJ, Gavin AT, Reid JS, O'Reilly D. Accuracy of hospital discharge data for cancer registration and epidemiological research in Northern Ireland. *Cancer Causes Control* 2000;11:899-905.
41. Tingulstad S, Halvorsen T, Norstein J, Hagen B, Skjeldestad FE. Completeness and accuracy of registration of ovarian cancer in the cancer registry of Norway. *Int J Cancer* 2002;98:907-11.
42. Brewster D, Stockon D, Harvey J, Mackay M. Reliability of cancer registration data in Scotland, 1997. *Eur J Cancer* 2002;38:414-7.

Summary

Assessment of data quality in disease registries

A perfect medical registry should contain a single complete accurate record for every registrable case occurring among the residents of the target population. In practice, this ideal can never be achieved. It is crucial to know the extent of inadequacy of data completeness and accuracy, as such data are increasingly used in clinical and epidemiological research and the validity

of the results depends on the quality of the data. Data quality in disease registries is studied routinely and in special studies. This article presents a classification and literature review of different methods used in the study of data quality in medical registries.

Katrin.Lang@ut.ee