Disease-associated reference intervals for twenty laboratory tests in patients with rheumatoid arthritis, Crohn's disease or ulcerative colitis

Gustav Mikkelsen [a,b,*], Børge Lillebo [c,d], Arild Faxvaag [e,f]

[a] Department of Clinical Chemistry, St Olav's Hospital, Trondheim University Hospital, Postboks 3250 Torgarden, 7006, Trondheim, Norway
[b] Department of Clinical and Molecular Medicine, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[c] Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Kirkegata 2, 7601, Levanger, Norway
[d] Clinic of Medicine and Rehabilitation, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway
[e] Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[f] Department of Rheumatology, St Olav's Hospital, Trondheim University Hospital, Postboks 3250 Torgarden, 7006, Trondheim, Norway

ARTICLE INFO

ABSTRACT

Background: Population based reference intervals are fundamental for interpreting results for quantitative laboratory tests. In patients with a specific chronic disorder, however, results of various tests may regularly be different than in healthy individuals. *Health-associated* reference intervals may therefore have limited value in such patients. Instead, *disease-associated* reference intervals may be useful, as they describe the results distribution in populations resembling the specific patients. Few disease-associated reference intervals are available in the literature. The aim of this study was to estimate reference intervals for common laboratory tests for patient populations with rheumatoid arthritis, Crohn's disease or ulcerative colitis without significant comorbidity, using a novel algorithm.
Material and methods: Laboratory test results and hospital discharge diagnoses were collected for relevant patients. An algorithm was developed to identify discharge diagnoses significantly associated with high or low results for specific tests. After excluding patients with such diagnoses, reference intervals were estimated, representing results distributions in patients with each of the specific chronic disorders, but without significant comorbidity.
Results: Disease-associated reference intervals were estimated for 20 common laboratory tests. Most of the estimated reference limits were significantly different from corresponding health-associated reference limits. Thirty percent of the estimated reference intervals were different from estimates based on crude patient populations, indicating that the algorithm applied managed to exclude patients with relevant comorbidity.
Conclusion: Disease-associated reference intervals could be estimated for a number of tests in patients with rheumatoid arthritis, ulcerative colitis or Crohn's disease using a highly automated algorithm based on routinely recorded patient data.

---

* Corresponding author. Department of Clinical Chemistry St Olavs hospital Postboks, 3250 Torgarden N, 7006, Trondheim, Norway.
E-mail addresses: gustav.mikkelsen@stolav.no (G. Mikkelsen), borge.lillebo@ntnu.no (B. Lillebo), arild.faxvaag@ntnu.no (A. Faxvaag).

1. Introduction

Population based reference intervals describe the expected result distribution for a test in a defined population [1]. They are essential for interpreting patient results and are often an integral part of test results reported from laboratories [2,3]. In order to be useful, reference intervals must be based on populations resembling the patient [2,3] and samples used for reference interval studies should be handled and analyzed in the same way as patient samples [4]. Hence, each laboratory should ensure that reference intervals are adequate for the population it serves and the analytical methods used [2,3,5,6].

In most cases, reference intervals reported by the laboratory describe the results expected in healthy populations, i.e. they are *health-associated* reference intervals [1,2,5]. They are useful to decide whether results from a specific patient are as expected in a *healthy* individual or not. Patient results outside the health-associated reference interval may indicate that the patient is not *healthy* and needs further evaluation. However, in patients with some chronic disorders, laboratory test results are frequently different than in healthy individuals due to the effects of the disorder on different organ systems and physiological functions. Results distributions of some tests may therefore be different in patients with specific chronic disorders than in healthy individuals. Using *health-associated* reference intervals when interpreting results from these patients may not be optimal, as a higher frequency of results outside these reference intervals may be expected than in healthy individuals. When interpreting test results for patients with chronic disorders, *disease-associated* reference intervals could be used [1,2,5,7–9], as they describe the results distribution in populations having the same disorder as the patient. Depending on the clinical question, such reference intervals may be a more relevant standard of reference than corresponding health-associated reference intervals.

Compared to the extensive literature on health-associated reference intervals, relatively few disease-associated reference intervals are available in the literature. Although data about patients with various chronic disorders are frequently available in many laboratories and health care institutions, using available information to establish disease-associated reference intervals is not frequently done. One reason for this may be the heterogeneity of many patient populations with respect to clinical picture, disease duration, comorbidity etc. Identifying test results from representative patients in order to establish reference intervals may therefore be challenging.

Laboratory tests play an important role in monitoring patients with many chronic inflammatory disorders. These disorders frequently affect multiple organs and physiological systems. Rheumatoid arthritis (RA) may impair the function of e.g. lungs, kidneys, hematopoiesis and cardiovascular system [10,11]. Inflammatory bowel disease (IBD) may cause renal, hepatobiliary, thromboembolic and nutritional complications [12–14]. Results for some laboratory tests may therefore be expected to be different in these patients compared to healthy individuals. The aim of this study was to establish disease-associated reference intervals for commonly ordered quantitative laboratory tests for patient populations with rheumatoid arthritis, Crohn's disease (CD) or ulcerative colitis (UC) without significant comorbidity, using routinely recorded patient data and a novel and highly automated algorithm.

2. Material and method

The setting was a university hospital with approximately 1000 beds. Data about test results reported by the clinical chemistry laboratory was collected for 20 different tests. The tests were chosen because they are commonly performed in many patients and the analytical methods have been relatively stable in the laboratory in the time period included. A summary of analytical methods is available in Supplemental Table 1. Patient demographics and all recorded discharge diagnoses for inpatient and outpatient encounters were collected for patients 18 years or older with at least one result for at least one of the included tests. Discharge diagnoses had been coded according to International Statistical Classification of Diseases and Related Health Problems - 10th revision (ICD-10).

For each included test, a *total population* (Fig. 1) was defined consisting of all patients having at least one result reported for the specific test between January 2005 and March 2018 and at least one ICD-10 code assigned between January 2000 and March 2018. Individual total populations were defined for males (M), for females (F) and for males and females together (B). The last result from each patient in each total population was included. A variable indicating whether specific ICD-10 codes had been assigned at least once to individual patients in each *total population* was constructed for each ICD-10 code occurring in the dataset. All ICD-10 codes were treated
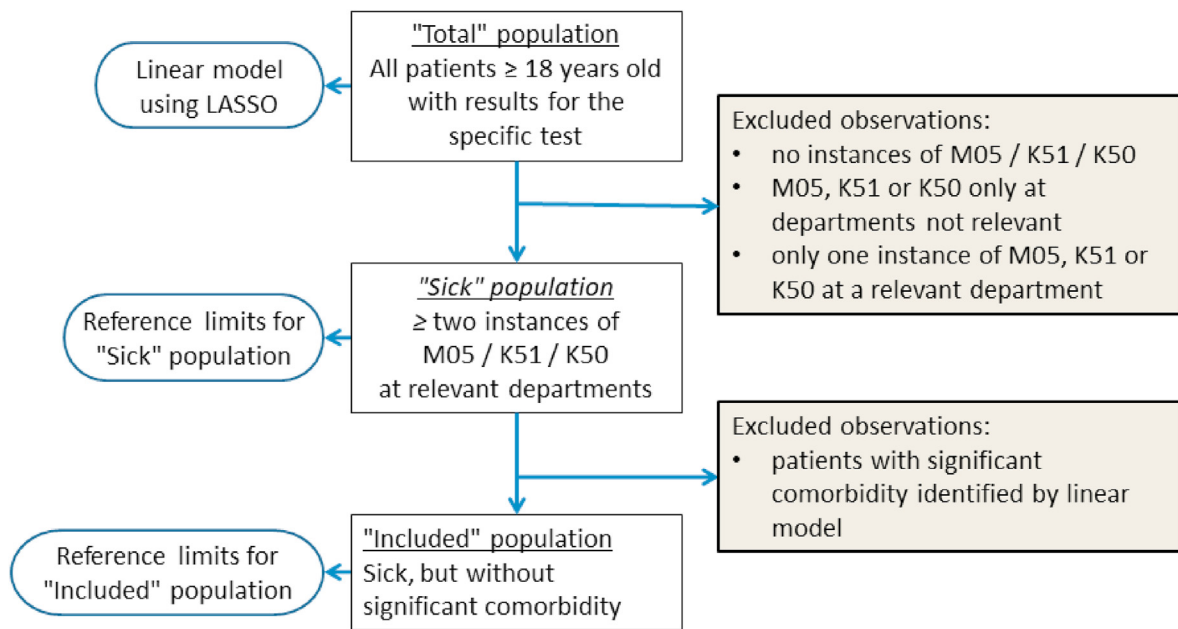
Fig. 1. Flowchart describing patient selection.

The diagram illustrates how populations used for estimating reference intervals were generated. A "total population" was identified for each of the 20 tests studied as patients having at least one result for the respective test. Individual *total* populations were defined for males and females separately and for males and females together, i.e. 180 populations in total for the twenty tests and three disorders. From each total population a "sick population" was identified as patients having the specific chronic disorder. Finally, an "included population" was identified as patients in the *sick* population not having significant comorbidity. The last test result from each patient in the *sick* populations and the *included* populations were used to estimate the respective reference intervals.

without suffixes. ICD-10 codes statistically associated with high or low result values of each specific test were identified in each total population using *least absolute shrinkage and selection operator* (LASSO) [15]. LASSO performs a least squares optimization, but with "shrinkage" of the coefficients. Small coefficients are given the value zero, causing the selected models to contain a relatively small number of predictor variables, which in our case were the individual indicator variables for ICD-10 codes assigned to patients. The number of predictors in the model derived by LASSO is influenced by the shrinkage coefficient, which was chosen for each model based on fivefold cross validation [15].

Patients having RA, CD or UC were identified in each *total population*. Based on experience with local coding practices, patients who had been assigned relevant ICD-10 codes, i.e. M05, K50 or K51 on at least two different hospital encounters at relevant hospital departments were considered to have RA, CD or UC, respectively. Using these criteria, one *sick population* was defined for each of the 20 tests for each of the three chronic disorders and each sex (Fig. 1). Patients in each *sick population* who had recorded any other ICD-10 codes identified by the LASSO selection procedure as significantly associated with high or low result values for the specific test were excluded. The remaining patients constituted patient populations likely to have the specific chronic disorder but unlikely to have comorbidity significantly associated with high or low results for the specific test. These patients were categorized as *included*. The results distributions in the *included* and *sick* populations were studied, and non-parametric reference limits were estimated as 2.5- and 97.5-percentiles with 90% bootstrap confidence intervals in populations with 80 or more results, subjectively considered to be an absolute minimum to estimate valid reference limits with acceptable precision. We also studied the association between estimated reference limits and patient age using quantile regression [16], and reference intervals were constructed for age groups when there was a statistically significant association of the 2.5- or the 97.5-percentile with age ($p < 0.05$). Partitioning according to age was done by constructing five year age groups and combining these with neighboring groups into partitions of minimum 80 results and with the 90% confidence interval of at least one reference limit not overlapping with the corresponding limit of at least one neighboring partition. No outlier values based on statistical principles were removed. Additional information about the algorithm is in the Supplement. All data analyses were done with SAS version 9.4 (SAS Institute, Cary, NC, USA).

The study was carried out in full accordance with the ethical principles of the Declaration of Helsinki and was approved by the Regional Committee for Medical and Health Research Ethics.

3. Results

3.1. Patients with RA, CD and UC

Approximately 18 million results were identified for the 20 tests included in this study, belonging to approximately 209 000 different patients, of which almost 200 000 had in total 6.4 million hospital inpatient and outpatient encounters between 2000 and 2018. 2477

patients had at least one instance of ICD-10 code M05 (RA), 1644 had at least one instance of K50 (CD) and 2842 had at least one instance of K51 (UC). 1602 patients fulfilled the criteria for RA, 320 for CD and 325 for UC as described above. Seventy-two percent, 54% and 42% of patients with RA, CD and UC were women, respectively. A few patients were identified as having more than one of the chronic disorders (Supplemental Table 2).

3.2. Results used for estimating reference limits

The number of patients in each of the 180 identified sick populations varied between 84 and 1600. Age varied between 18 and 95 years and median age was 68, 35 and 38 years for patients identified as having RA, CD and UC, respectively (Fig. 2). The time between individual patients were diagnosed with the specific chronic disorder and sample collection varied between 0.3 and 18 years (Fig. 3), and median (1st - 3rd quartile) time since diagnosis was 7.8 (4.1–12.8), 7.6 (3.8–12.2) and 6.3 (3.1–10.6) years for RA, CD and UC, respectively.

3.3. Reference intervals

Estimated reference intervals for *sick* and *included* populations consisting of more than 80 patients with CD, RA or UC but without significant comorbidity, are presented in Tables 1–3, respectively. When confidence intervals of estimated limits overlap with corresponding limits for the other sex we suggest using common reference intervals for males and females. For 30% of the suggested reference intervals the confidence intervals of upper or lower reference limits for the *included* populations did not overlap with the corresponding limit for the *sick population*. Confidence intervals of most limits for *included* populations did not overlap with corresponding health-associated reference limits reported by the laboratory, although differences were in some cases moderate.

For 11 populations there was a statistically significant association between the upper or lower reference limit and patient age, as illustrated in Supplemental Fig. 1. Reference limits for age groups and quantile regression estimates for continuous, age dependent reference limits are presented in Supplemental Table 4 and Supplemental Table 5, respectively.
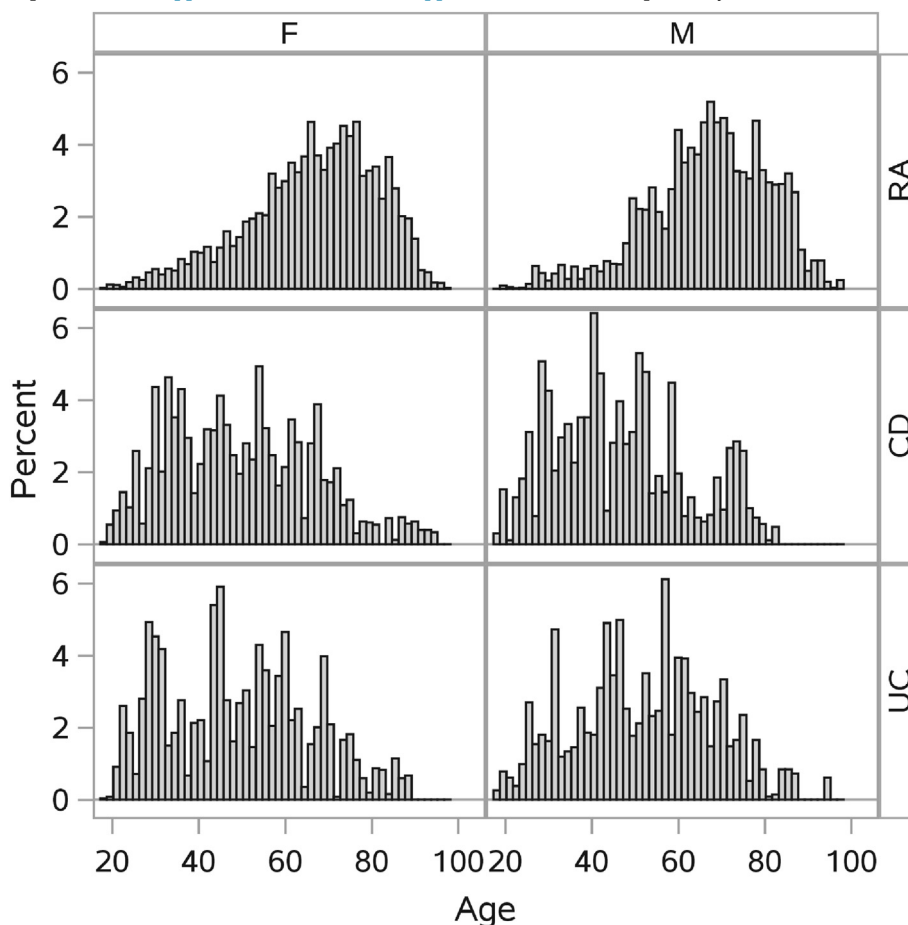

Fig. 2. Age distributions.
Histograms illustrate age distributions for all patients identified as having RA, CD or UC.
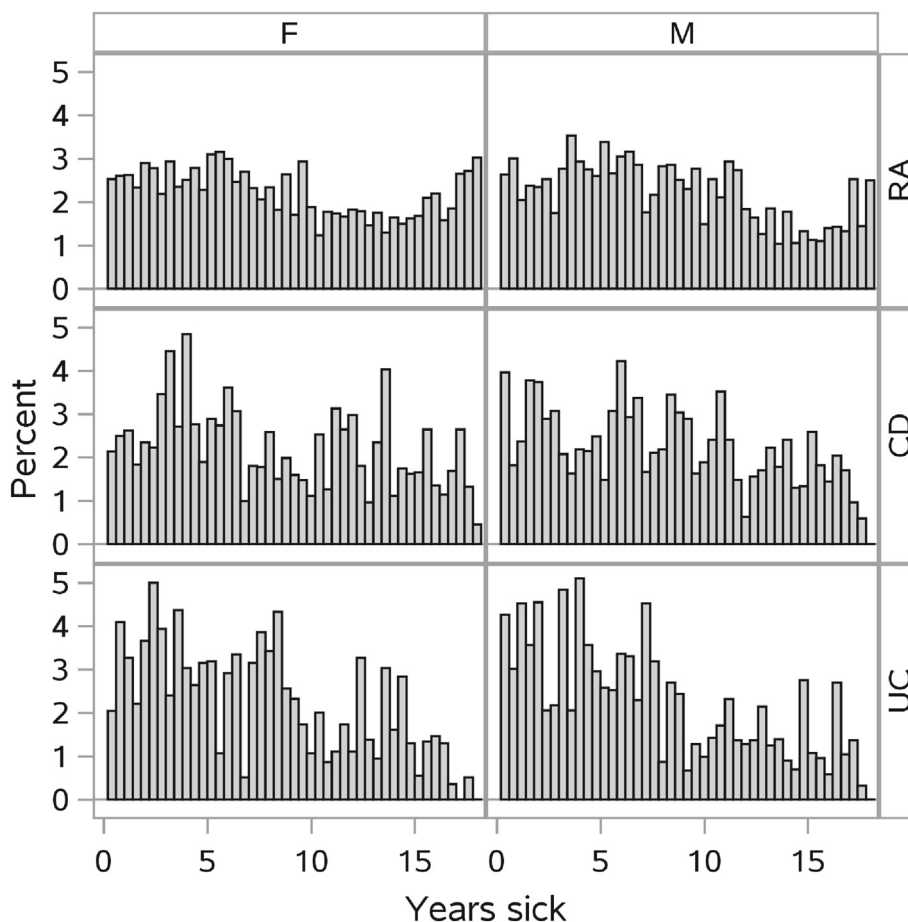RA: rheumatoid arthritis; CD: Crohn's disease; UC: ulcerative colitis.

Fig. 3. Time since diagnosis.
Histograms illustrate distributions of time between diagnosis and sample collection for each result used for estimating reference intervals.
RA: rheumatoid arthritis; CD: Crohn's disease; UC: ulcerative colitis.

3.4. Identified models

Number of ICD-10 codes identified as significantly associated with the results of the individual tests in each included population varied between 2 and 86 (median 8) as illustrated in Supplemental Fig. 2. Numbers of parameters were in general higher in populations excluded from further analysis due to small number of remaining results. Considering that there should in general be at least 10 observations for each parameter in the model [17], a few of the models used to identify populations for which reference limits were estimated may have been overfitted (Supplemental Fig. 3), leading to small sample sizes.

4. Discussion

The idea of *disease-associated* reference intervals is not new [1,2,5,7–9,18–20], but this is the first published study using LASSO selection of discharge diagnoses to identify patients without comorbidity significantly associated with high or low result values for specific laboratory tests in a general hospital population. The suggested reference intervals may be more representative for patients with RA, CD or UC than reference intervals based on healthy populations or on crude patient populations from which patients with relevant comorbidity have not been excluded. The suggested reference limits may be useful when interpreting results for patients with these chronic disorders, because they describe which *concrete* result values to expect for commonly performed laboratory tests in patients with RA, CD or UC, respectively, and without additional disorders significantly interfering with the results. Patients with RA, CD or UC with results outside the respective disease associated reference intervals may require further evaluation. Although we chose to study RA, CD and UC, the algorithm applied may be useful to estimate reference intervals for other disorders and in other settings.

4.1. Results distributions

As expected, markers of inflammation tended to be increased in all the studied populations compared to values expected in healthy individuals (Tables 1–3). Markers like C-reactive protein (CRP), leukocyte count and ESR are influenced by e.g. intercurrent infections. If

Table 1
Suggested reference limits for Crohn's disease.

Test Unit	Sex	Included populations						Sick populations					Health-associated reference interval
		N	%	Lower	90%CI	Upper	90%CI	N	Lower	90%CI	Upper	90%CI	
ALP U/L	F	158	92%	**41**	(36–44)	**139**	(129–164)	172	41	(36–44)	142	(136–220)	*35–105*
ALT U/L	F	165	96%	**9**	(7-9)	**63**	(43–79)	172	9	(6-9)	63	(51–79)	*10–45*
Bilirubin μmol/L	B	251	81%	**3**	3,4(3-4)	**27**	(21-28)	311	3	(3-4)	28	(23–38)	*5–25*
CK U/L	F	87	100%	**23**	(10-31)	**265**	(139–498)	87	23	(10-31)	265	(139–498)	*35–210*
Creatinine μmol/L	M	135	94%	**51**	(50–60)	**133**	(110–165)	144	51	(50–60)	*175*	*(133–192)*	*60–105*
	F	142	83%	**39**	(38–43)	**101**	(83–103)	171	39	(38–43)	*115*	*(101–140)*	*45–90*
Ferritin μg/L	M	138	99%	**15**	(7-18)	**448**	(355–1158)	139	15	(7-18)	448	(355–1158)	*30–383*
	F	169	99%	**13**	(7-15)	**319**	(262–1045)	171	13	(7-15)	319	(262–1045)	*20–167*
GT U/L	B	223	72%	**10**	(10-10)	**185**	(117–227)	309	10	(10-10)	210	(140–308)	*F < 40: 10–45; F ≥ 40: 10–75; M < 40: 10–80; M ≥ 40: 15-115*
IgG g/L	F	81	96%	**5.3**	(5.1–6.5)	**17.1**	(15.7–19.8)	84	5.1	(4.1–6.5)	17.1	(15.5–19.8)	*19-50: 6.9-15.7; > 50: 6.1-14.9*
Leukocytes x 10^9/L	B	202	64%	**3.8**	(3.5–4.1)	**12.1**	(11.3–14.5)	317	3.8	(3.5–4.2)	*16.1*	*(12.2-20.9)*	*4.1–9.8*
Platelets x 10^9/L	B	81	26%	**141**	(116–179)	**468**	(412–544)	316	133	(116–148)	499	(451–544)	*164–370*
TSH mIU/L	M	111	100%	**0.48**	(0.01–0.60)	**4.12**	(3.44–4.87)	111	0.48	(0.01–0.60)	4.12	(3.44–4.87)	*0.24–3.78*
	F	149	96%	**0.04**	(0.02–0.21)	**3.97**	(3.49–7.04)	156	0.04	(0.02–0.21)	4.01	(3.61–9.79)	*0.24–3.78*
eGFR ml/min/1.73m²	M	80	56%	**60**	(49–75)	**137**	(123–138)	143	*33*	*(29–46)*	131	(125–138)	*> 60*

Table presents estimated reference limits for patients with Crohn's disease for identified populations with more than 80 patients. Italic text indicates reference limits for which 90% confidence intervals of corresponding limits for *included* populations do not overlap. Age distributions for individual populations are available in Supplemental Table 3.

Sick population: all patients identified as having RA, CD or UC, respectively. Included population: patients identified as having RA, CD or UC but no significant comorbidity. N: number of observations. %: fraction of *sick* population in the *included* population. Lower: the 2.5th percentile in the results distribution. Upper: the 97.5th percentile in the results distribution. 90%CI: 90% confidence interval. ALP: alkaline phosphatase. ALT: alanine aminotransferase. CK: creatine kinase. CRP: C-reactive protein. ESR: erythrocyte sedimentation rate. GT: gamma-glutamyltransferase. IgG: immunoglobulin G. TSH: thyroid stimulating hormone. eGFR: estimated glomerular filtration rate.

Table 2
Suggested reference limits for rheumatoid arthritis.

Test	Sex	Included populations						Sick populations					Health-associated reference interval
		N	%	Lower	90%CI	Upper	90%CI	N	Lower	90%CI	Upper	90%CI	
ALP U/L	B	1165	89%	**37**	(36–39)	**181**	(158–200)	1306	37	(36–39)	*243*	*(208–293)*	*35–105*
ALT U/L	F	1137	99%	**8**	(7-9)	**74**	(66–87)	1145	8	(7-9)	76	(70–91)	*10–45*
Albumin g/L	F	101	13%	**33**	(20–36)	**47**	(47–49)	771	21	(21-23)	*46*	*(46–47)*	18-39: 36–48; 40–70: 36–45; >70: 34-45
Bilirubin µmol/L	M	248	89%	**3**	(3-4)	**29**	(21–36)	279	3	(3-4)	32	(23–55)	*5–25*
	F	528	80%	**3**	(3-4)	**21**	(20-24)	662	3	(3-3)	24	(21-28)	*5–25*
CK U/L	M	261	99%	**21**	(17–25)	**386**	(289–736)	264	21	17-25)	386	(289–736)	*< 50: 50–400; ≥ 50: 40-280*
	F	680	100%	**17**	(15-20)	**261**	(210–326)	682	17	(15-20)	259	(205–326)	*35–210*
CRP mg/L	M	143	32%	**5**	(5-5)	**31**	(19–42)	450	5	(5-5)	*209*	*(164–228)*	*< 5*
	F	213	19%	**5**	(5-5)	**114**	(30–127)	1140	5	(5-5)	*199*	*(173–221)*	*< 5*
Calcium mmol/L	F	269	36%	**2.09**	(2.06–2.14)	**2.55**	(2.51–2.55)	748	*1.99*	*(1.91-2.02)*	2.57	(2.55-2.63)	*2.15–2.51*
Cholesterol mmol/L	F	85	13%	**3.8**	(3.4–4.2)	**7.5**	(7.0–8.0)	679	3	(2.9-3.3)	7.7	(7.4–8.0)	*<30: 2.9–6.1; 30–49: 3.3–6.9; ≥50: 3.9–7.8*
Creatinine µmol/L	M	192	43%	**56**	(53–62)	**105**	(98–119)	451	55	(51–58)	*187*	*(162–205)*	*60–105*
	F	739	64%	**41**	(39–42)	**93**	(87–95)	1149	39	(37–41)	*145*	*(130–177)*	*45–90*
ESR mm/h	B	219	18%	**2**	(1-2)	**60**	(44–78)	1196	2	(2-2)	*92*	*(85–96)*	*F < 45: ≤ 14; F > 45: ≤ 24; M < 45: ≤ 7; M > 45: ≤ 17*
Ferritin µg/L	F	889	99%	**15**	(14-16)	**700**	(563–822)	896	15	(14-16)	714	(589–867)	*20–167*
Folate nmol/L	F	290	48%	**6**	(4-7)	**54**	(54-54)	610	6	(5-6)	54	(54-54)	*7–29*
GT U/L	M	250	77%	**11**	(10-12)	**250**	(202–447)	324	11	(10-13)	*607*	*(247–798)*	*< 40: 10–80; ≥ 40: 15-115*
	F	661	82%	**10**	(10-10)	**170**	(141–221)	811	10	(10-10)	*299*	*(260–387)*	*< 40: 10–45; ≥ 40: 10-75*
Hemoglobin g/dL	M	83	18%	**13.2**	(12.7–13.5)	**16.6**	(16.3–16.8)	449	*9.7*	*(9.2-10.0)*	16.7	(16.4–16.9)	*13.4–17*
	F	206	18%	**10.8**	(10.3–11.5)	**15.2**	(15.0–15.7)	1145	*9.4*	*(9.1-9.8)*	15.6	(15.3–15.7)	*11.7–15.3*
IgG g/L	M	174	99%	**5.1**	(3.8–5.9)	**19.8**	(17.2–21.5)	175	5.1	(3.8–5.9)	19.8	(17.2–21.5)	*6.1–14.9*
	F	402	78%	**5.1**	(4.8–5.7)	**17.3**	(16.0–17.7)	513	4.8	(4.6–5.2)	17.6	(16.6–18.2)	*19-50: 6.9-15.7; > 50:6.1-14.9*
Leukocytes x 10^9/L	M	444	99%	**3.7**	(3.5–4.1)	**18.4**	(15.7–23.4)	449	3.7	(3.1–4.0)	18.6	(15.7–27.1)	*4.1–9.8*
	F	777	68%	**3.9**	(3.8–4.0)	**14.2**	(13.6–14.5)	1146	3.8	(3.7–4.0)	*17.4*	*(16.5-19.4)*	*4.1–9.8*
Platelets x 10^9/L	M	131	29%	**144**	(117–164)	**374**	(351–412)	448	*98*	*(80–135)*	*441*	*(405–544)*	*164–370*
	F	979	86%	**156**	(147–165)	**486**	(461–536)	1139	*140*	*(121–151)*	495	(464–536)	*164–370*
Sodium mmol/L	F	474	52%	**134**	(133–135)	**145**	(145–146)	912	133	(131–133)	145	(145–146)	*137–145*
TSH mIU/L	M	320	98%	**0.48**	(0.42–0.63)	**4.93**	(4.54–5.48)	326	0.46	(0.34–0.61)	4.93	(4.54–5.72)	*0.24–3.78*
	F	852	91%	**0.14**	(0.07–0.25)	**5.52**	(4.83–6.36)	937	0.07	(0.02–0.11)	6.1	(5.21–6.63)	*0.24–3.78*
eGFR ml/min/1.73m^2	M	131	35%	**67**	(54–69)	**117**	(113–123)	374	*40*	*(35–45)*	117	(113–119)	*> 60*
	F	232	25%	**56**	(49–64)	**116**	(113–122)	930	*35*	*(28–42)*	*123*	*(120–124)*	*> 60*

Table presents estimated reference limits for patients with rheumatoid arthritis. See footnote for Table 1 for explanations.

Table 3
Suggested reference limits for ulcerative colitis.

Test	Sex	N	%	Lower	90%CI	Upper	90%CI	N	Lower	90%CI	Upper	90%CI	Health-associated reference interval
			Included populations					Sick populations					
ALP U/L	M	128	69%	**44**	(31–46)	**160**	(141–212)	185	41	(30–45)	*306*	*(215–477)*	*35–105*
	F	108	81%	**31**	(25–36)	**134**	(123–175)	134	33	(30–37)	*393*	*(146–458)*	*35–105*
ALT U/L	F	133	98%	**9**	(7-10)	**85**	(52–162)	136	9	(7-10)	85	(58–162)	*10–45*
Bilirubin μmol/L	M	147	82%	**3**	(3-4)	**43**	(26–49)	179	4	(3-4)	49	(32–62)	*5–25*
	F	114	88%	**3**	(3-4)	**25**	(19–32)	130	3	(3-4)	26	(20–59)	*5–25*
CK U/L	M	85	100%	**20**	(15-29)	**639**	(337–5573)	85	20	(15-29)	639	(337–5573)	*< 50: 50–400; ≥ 50: 40-280*
CRP mg/L	B	86	27%	**5**	(5-5)	**61**	(43–121)	318	5	(5-5)	114	(72–157)	*< 5*
Calcium mmol/L	B	82	28%	**2.17**	(2.15–2.21)	**2.56**	(2.49–2.59)	296	2.04	(1.91-2.11)	2.56	(2.54–2.58)	2.15–2.51
Creatinine μmol/L	M	169	90%	**53**	(44–59)	**134**	(104–183)	188	53	(44–60)	*216*	*(161–363)*	60–105
	F	100	74%	**44**	(43–49)	**88**	(82–98)	136	44	(43–49)	*125*	*(89–138)*	45–90
Ferritin μg/L	F	126	98%	**8**	[5–11]	**284**	(224–358)	129	8	[5–11]	308	(225–358)	*20–167*
GT U/L	M	138	75%	**11**	(10-12)	**202**	(125–402)	183	11	(10-14)	*557*	*(334–1690)*	*< 40: 10–80; ≥ 40: 15-115*
	F	105	80%	**9**	(9-10)	**167**	(100–594)	131	10	(9-10)	333	(146–484)	*< 40: 10–45; ≥ 40: 10-75*
IgG g/L	B	104	94%	**6.8**	(5.7–7.3)	**22.4**	(18.7–28.0)	111	6.8	(5.7–7.3)	22.4	(18.7–28.0)	*F:19–50: 6.9-15.7; F > 50: 6.1-14.9; M: 6.1-14.9*
Leukocytes x 10^9/L	M	143	76%	**4.2**	(3.7–4.5)	**14.8**	(12.9–17.0)	187	4.0	(3.6–4.3)	*17.4*	*(14.5-19.6)*	*4.1–9.8*
	F	110	81%	**3.7**	(3.1–4.0)	**12.8**	(11.2–16.8)	136	3.8	(3.4–4.0)	12.4	(11.2–13.9)	*4.1–9.8*
TSH mIU/L	M	140	95%	**0.39**	(0.25–0.52)	**5.21**	(4.37–9.22)	147	0.36	(0.11–0.50)	5.21	(4.37–9.22)	*0.24–3.78*
	F	108	92%	**0.23**	(0.01–0.51)	**3.93**	(3.35–4.19)	118	0.19	(0.01–0.40)	3.98	(3.50–4.89)	0.24–3.78

Table presents estimated reference limits for patients with ulcerative colitis. See footnote for Table 1 for explanations.

treated outside the hospital such disorders are not always reflected in the discharge diagnoses recorded by the hospital. It is therefore possible that estimated reference intervals for these tests are higher than in populations with RA, CD or UC without any individuals with intercurrent infections. The suggested reference intervals may still be more representative for these patients than corresponding intervals based on healthy populations or populations with RA, CD or UC from which no patients with other relevant comorbidities have been excluded.

Reference intervals for ferritin in patients with RA, CD and UC were wider than reference limits for healthy populations (Tables 1–3). Lower limits for females were lower for patients with UC compared to RA and CD, and upper limits were higher in patients with RA compared to UC and CD. This is compatible with that patients with UC more often have iron deficiency due to chronic blood loss and malabsorption [13,14,21,22]. Patients with RA appear to have more prominent inflammation based on results for leukocyte counts and CRP, possibly explaining higher levels of ferritin.

In patients with RA, estimates for the lower reference limits for hemoglobin in the *included* populations were only slightly lower than the corresponding health-associated reference limits reported by the laboratory (Table 2). Low hemoglobin concentrations were thus not a very prominent characteristic of patients in this population. This is somewhat unexpected, as RA generally is associated with anemia of chronic disease [10,23]. A possible explanation could be that many of the patients studied receive adequate treatment in our setting and therefore do not develop prominent anemia.

The lower reference limits for creatine kinase (CK) in *included* patients with RA were somewhat lower than the corresponding health associated reference limits. As CK correlates with muscle mass [24] and patients with RA are at risk of developing muscle atrophy [25, 26], low concentrations of CK are not unexpected. The estimated reference limits indicate how much lower concentrations to expect.

4.2. Disease-associated reference intervals

The literature describing the characteristics of patients with RA, CD or UC is extensive. However, it often lacks specific and systematic descriptions of what concrete result values to expect for frequently ordered laboratory tests in these patients [10–14,23], as opposed to the voluminous literature on reference intervals for healthy populations. Specific reference information to be used when interpreting test results for patients with these chronic disorders may therefore be limited. Even if available, published descriptions may not be valid in local settings as laboratory methods, patient demographics, ethnicity, disease severity and therapeutic strategies may differ substantially between health care institutions and geographical regions. Reference intervals should therefore be derived locally [2, 3].

Disease-associated reference intervals could be a valuable supplement to the health-associated reference intervals routinely reported by clinical laboratories. We believe that disease-associated reference limits have a potential to improve interpretation of laboratory test results in clinical settings, as they represent systematic descriptions of which result values to expect in populations frequently receiving health care. Instead of relying on personal experience, locally derived reference limits for relevant patient populations can make this knowledge available for any clinician, including those with less experience in handling patients with the specific chronic disorders. In addition, such reference intervals may be estimated on basis of larger number of subjects than any clinician have encountered and may therefore be a valuable source of evidence based, objective knowledge for any one treating these patients. As data needed to establish such estimates are often available, reference intervals for specific patient populations can be estimated without the costs of sample collection and analysis associated with establishing population based health-associated reference intervals. Estimating reference limits based on available data can be done with low cost and estimates can easily be updated on a regular basis. In time, as more data becomes available, estimates are likely to become more representative and precise.

Disease associated reference intervals derived in a consistent way may not only be useful when interpreting results for individual patients. They can also be used to describe characteristics of patient populations and hence serve as basis for decision support systems. By comparing reference intervals derived systematically for corresponding populations from different settings they can also be used to compare characteristics of patient populations receiving different health care and follow-up, and hence serve as quantitative quality indicators.

4.3. Algorithms used

In general, estimated reference intervals tended to be narrower in the *included* than in the *sick* populations. This indicates that the algorithm did manage to remove patients with comorbidity associated with high or low result values of specific tests based on routinely recorded ICD 10-codes. In some populations, however, almost no patients were removed by the algorithm. This may either be due to failure of the algorithm to identify significant comorbidity or that no patients with significant comorbidity existed. In other populations, for which estimates are not presented in Tables 1–3, a high number of patients were removed. This indicates that the algorithm failed to identify a sufficient number of reference individuals without significant comorbidity, possibly because too few such patients existed in the population.

The algorithm applied is highly automated. Selection of patients is not influenced by extensive subjective assessments, as is frequently an issue when patient populations are identified manually. The algorithm is highly flexible and can easily be extended with other types of information to identify even more representative populations. Adjustments can be done to optimize individual models, using subjective or objective criteria to extend or simplify identified models. We did not examine each linear model identified by the LASSO selection process in detail. Thus, we did not try to explain *why* specific patients were excluded or included in the populations identified. Doing so might have revealed information that could be used to refine the models applied, but this was outside the scope of this study.

No outlier values based on statistical principles were removed. As patients with RA, CD or UC by nature are likely to be more heterogeneous than healthy individuals, we believe that excluding values based on how they fit in the results distributions would obscure some of the characteristics of these patient populations [27].

4.4. Limitations of the study

Although the volume of data available in our study was significant, small populations was a limitation for estimating reference intervals for some tests. Hence, reference intervals are suggested for only some of the studied tests and patient groups (Tables 1–3). For the same reason, partitioning into age groups in order to estimate age dependent reference limits had to be done in a pragmatic way to obtain age groups of adequate size and reference limit estimates reflecting the association between test results and age in a reasonable way.

As test results were included from a long time period, different analytical assays were used for some of the tests (Supplemental Table 1). Although we only included tests for which the laboratory had used analytical methods with similar characteristics in the relevant time period, many populations have results from more than one analytical method. Although this can make it more difficult to interpret the estimates, it can also make them more generalizable, as between method variations are included.

We did not verify manually that individual patients identified as having RA, CD or UC actually suffered from the disorders. We also did not validate the discharge diagnoses used to identify patients with significant comorbidity, but relied on the records made by treating clinicians. To validate discharge diagnoses would require a manual evaluation of thousands of patients, and this was beyond our capabilities. Some patients may therefore have been incorrectly included in or excluded from the reference sample groups, as is a general problem when identifying reference individuals in any setting. Based on experience with local coding practices and that the numbers of patients identified were approximately as expected [28–31], we feel confident that a high proportion of the patients identified as having either RA, CD or UC actually suffered from the respective disorders and that they represent significant fractions of the relevant hospital populations. The risk of incorrectly excluding patients actually having one of the chronic disorders was probably higher than the risk of including patients without the disorder. This should however have moderate effects on the estimated reference intervals as long as the patients included are still representative for the populations studied. Although the reference sample groups identified are not perfect we believe that they are a more relevant standard of reference than healthy populations or crude patient populations from which patients with other relevant comorbidities have not been removed.

In conclusion, we have estimated disease associated reference intervals for several commonly performed tests for patients with RA, CD or UC, respectively, by developing and applying a novel and highly automated algorithm in an unselected patient population to identify relevant patients and eliminate subjects with significant comorbidity. As few disease-associated reference intervals are available in the literature, our estimates may be useful for interpreting test results for patients with the disorders studied. The approach described may be applied to estimate reference intervals for other patient populations and other laboratory tests in different settings.

CRediT authorship contribution statement

Gustav Mikkelsen: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing – original draft. **Børge Lillebo:** Conceptualization, Methodology, Writing – review & editing. **Arild Faxvaag:** Conceptualization, Methodology, Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plabm.2021.e00225.

References

[1] H.E. Solberg, International federation of clinical chemistry (IFCC), scientific committee, clinical section, expert panel on theory of reference values, and international committee for standardization in haematology (ICSH), standing committee on reference values. Approved recommendation (1986) on the theory of reference values. Part 1. The concept of reference values, J. Clin. Chem. Clin. Biochem. 25 (5) (1987) 337–342.

[2] G. Horowitz, G.R.D. Jones, Establishment and Use of Reference Intervals, in: N. Rifai, A.R. Horvath, C.T. Wittwer (Eds.), Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, sixth ed., Elsevier Saunders, 2018, pp. 170–194. ISBN: 978-0-323-35921-4.

[3] K. Adeli, F. Ceriotti, M. Nieuwesteeg, Reference Information for the Clinical Laboratory, in: C.A. Burtis, E.R. Ashwood, D.E. Bruns (Eds.), Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, sixth ed., Elsevier Saunders, 2018, pp. 1745–1818. ISBN: 978-0-323-35921-4.

[4] H.E. Solberg, C. PetitClerc, International federation of clinical chemistry (IFCC), scientific committee, clinical section, expert panel on theory of reference values. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values, J. Clin. Chem. Clin. Biochem. 26 (9) (1988) 593–598.

[5] C.L.S.I. Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline, in: third ed.CLSI Document C28-A3c, 2008, p. 37.

[6] E.K. Harris, J.C. Boyd, Statistical Bases of Reference Values in Laboratory Medicine, M. Dekker, 1995, p. 361. ISBN: 0-8247-9339-0.

[7] H.E. Solberg, Subject-based reference values, Scand. J. Clin. Lab. Invest. Suppl. 222 (1995) 7–10, https://doi.org/10.3109/00365519509088444.

[8] B. Dugue, E. Leppanen, R. Grasbeck, Preanalytical factors and the measurement of cytokines in human subjects, Int. J. Clin. Lab. Res. 26 (2) (1996) 99–105, https://doi.org/10.1007/bf02592351.

[9] E.M. Wright, P. Royston, Calculating reference intervals for laboratory measurements, Stat. Methods Med. Res. 8 (2) (1999) 93–112, https://doi.org/10.1177/096228029900800202.

[10] R.A. Watts, P.G. Conaghan, C. Denton, H. Foster, J. Isaacs, U. Müller-Ladner, Oxford Textbook of Rheumatology, 4 ed, Oxford University Press, 2013. ISBN: 9780199642489.

[11] J.S. Smolen, D. Aletaha, I.B. McInnes, Rheumatoid arthritis, Lancet 388 (10055) (2016) 2023–2038, https://doi.org/10.1016/S0140-6736(16)30173-8.

[12] D.C. Baumgart, W.J. Sandborn, Crohn's disease, Lancet 380 (9853) (2012) 1590–1605, https://doi.org/10.1016/S0140-6736(12)60026-9.

[13] R. Ungaro, S. Mehandru, P.B. Allen, L. Peyrin-Biroulet, J.F. Colombel, Ulcerative colitis, Lancet 389 (10080) (2017) 1756–1770, https://doi.org/10.1016/S0140-6736(16)32126-2.

[14] WF Stenson, SB Hanauer, RD Cohen, Inflammatory Bowel Disease (2009) 1386–1472.

[15] T.J. Hastie, R.J. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2001.

[16] R. Koenker, Quantile Regression, Cambridge University Press, 2005. ISBN: 9780521845731.

[17] D.G. Altman, Practical Statistics for Medical Research, Chapman & Hall, London), 1991, p. 611.

[18] M. Lindberg, A. Hole, H. Johnsen, A. Åsberg, A. Rydning, H.E. Myrvold, et al., Reference intervals for procalcitonin and C-reactive protein after major abdominal surgery, Scand. J. Clin. Lab. Invest. 62 (3) (2002) 189–194.

[19] G. Mikkelsen, A. Åsberg, M.E. Hultström, K. Aasarød, G.G. Hov, Reference limits for chromogranin A, CYFRA 21-1, CA 125, CA 19-9 and carcinoembryonic antigen in patients with chronic kidney disease, Int. J. Biol. Markers 32 (4) (2017) e461–e466, https://doi.org/10.5301/ijbm.5000278.

[20] D. Monneret, P. Hausfater, B. Riou, D. Bonnefont-Rousselot, Estimated GFR-specific 99th percentiles for high-sensitive cardiac troponin T based on the adjusted analytical change limit (adjACL) in hospitalized patients, Clin. Chem. Lab. Med. 57 (2) (2018) e5–e7, https://doi.org/10.1515/cclm-2018-0485.

[21] F. Gomollon, J.P. Gisbert, Anemia and inflammatory bowel diseases, World J. Gastroenterol. 15 (37) (2009) 4659–4665, https://doi.org/10.3748/wjg.15.4659.

[22] P. Bager, R. Befrits, O. Wikman, S. Lindgren, B. Moum, H. Hjortswang, et al., The prevalence of anemia and iron deficiency in IBD outpatients in Scandinavia, Scand. J. Gastroenterol. 46 (3) (2011) 304–309, https://doi.org/10.3109/00365521.2010.533382.

[23] A. Wilson, H.T. Yu, L.T. Goodnough, A.R. Nissenson, Prevalence and outcomes of anemia in rheumatoid arthritis: a systematic review of the literature, Am. J. Med. 116 (Suppl 7A) (2004), https://doi.org/10.1016/j.amjmed.2003.12.012, 50S-7S.

[24] M. Panteghini, R. Bais, Serum Enzymes, in: C.A. Burtis, E.R. Ashwood, D.E. Bruns (Eds.), Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, sixth ed., Elsevier Saunders, 2018, pp. 404–434. ISBN: 9780323359214.

[25] R. Khurana, S.M. Berney, Clinical aspects of rheumatoid arthritis, Pathophysiology 12 (3) (2005) 153–165, https://doi.org/10.1016/j.pathophys.2005.07.009.

[26] G.J. Herbison, J.F. Ditunno, M.M. Jaweed, Muscle atrophy in rheumatoid arthritis, J. Rheumatol. Suppl. 14 (Suppl 15) (1987) 78–81.

[27] H.E. Solberg, Approved recommendation (1988) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits, Clin. Chim. Acta 170 (1987) S13–S32.

[28] D.L. Scott, F. Wolfe, T.W. Huizinga, Rheumatoid arthritis, Lancet 376 (9746) (2010) 1094–1108, https://doi.org/10.1016/S0140-6736(10)60826-4.

[29] V. Videm, R. Thomas, M.A. Brown, M. Hoff, Self-reported diagnosis of rheumatoid arthritis or ankylosing spondylitis has low accuracy: data from the Nord-Trøndelag health study, J. Rheumatol. 44 (8) (2017) 1134–1141, https://doi.org/10.3899/jrheum.161396.

[30] M.D. Kappelman, S.L. Rifas-Shiman, K. Kleinman, D. Ollendorf, A. Bousvaros, R.J. Grand, et al., The prevalence and geographic distribution of Crohn's disease and ulcerative colitis in the United States, Clin. Gastroenterol. Hepatol. 5 (12) (2007) 1424–1429, https://doi.org/10.1016/j.cgh.2007.07.012.

[31] M.H. Vatn, A.K. Sandvik, Inflammatory bowel disease, Scand. J. Gastroenterol. 50 (6) (2015) 748–762, https://doi.org/10.3109/00365521.2015.1033000.