BMJ Open

"OMICS" BIOMARKERS ASSOCIATED WITH CHRONIC LOW BACK PAIN: A RETROSPECTIVE CLINICAL STUDY

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012070
Article Type:	Protocol
Date Submitted by the Author:	30-Mar-2016
Complete List of Authors:	Allegri, Massimo; Universita degli Studi di Parma, Department of surgical science; Azienda Ospedaliera Universitaria Parma, Anesthesia Intensive Care and Pain Therapy service MINELLA, CRISTINA; Fondazione IRCCS Policlinico San Matteo, Anesthesia, Intensive Care and Pain Therapy, Emergency Department DE GREGORI, MANUELA; Fondazione IRCCS Policlinico San Matteo, Anesthesia, Intensive Care and Pain Therapy, Emergency Department Klersy, Catherine; Fondazione IRCCS Policlinico San Matteo, Service of Biometry & Satistics, Research Department Wang, Wei; Edith Cowan University, School of Medical and Health Sciences Sim, Moira; Edith Cowan University, School of Medical and Health Sciences Gieger, Christian; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit of Molecular Epidemiology; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Institute of Epidemiology II Manz, Judith; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit of Molecular Epidemiology; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Institute of Epidemiology II Pemberton, Iain; Photeomix, IP Research Consulting SAS MacDougall, Jane; Photeomix, IP Research Consulting SAS Williams, Frances; Kings College London, Twin Research and Genetic Epidemiology Depart Van Zundert, Jan; Ziekenhuis Oost-Limburg, Department of Anesthesiology, Critical Care and Multidisciplinary Pain Center Buyse, Klaas; Ziekenhuis Oost-Limburg, Department of Anesthesiology, Critical Care and Multidisciplinary Pain Center Gudelj, Ivan; Genos Glycoscience Research Laboratory Lauc, Gordan; Genos Glycoscience Research Laboratory Primorac, Dragan; St. Catherine Specialty Hospital; Sveuciliste u Splitu Medicinski fakultet Skelin, Andrea; Genos Glycoscience Research Laboratory; St. Catherine Specialty Hospital Aulchenko, Yurii; PolyOmica Karssen, Lennart; PolyOmica Kapural, Leonardo; Carolinas Pain Institute Rauck, Richard; Carolinas Pain Institute Fanelli, Guido; Universita degli Studi di Parma, Department of surgical science; Azienda Ospedaliero Universitaria di Parma, Anesthesia Intensive Care and Pain Therapy service

Primary Subject Heading:	Patient-centred medicine
Secondary Subject Heading:	Genetics and genomics, Research methods
Keywords:	Pain management < ANAESTHETICS, CHRONIC LOW BACK PAIN, GENOME-WIDE ASSOCIATION STUDY, GLYCOMICS, ACTIVOMICS

"OMICS" BIOMARKERS ASSOCIATED WITH CHRONIC LOW BACK PAIN: A RETROSPECTIVE CLINICAL STUDY

Authors: Massimo Allegri[1,2], Cristina E. Minella[3], Manuela De Gregori[3], Catherine Klersy[4], Wei Wang[5], Moira Sim[5], Christian Gieger[6,7], Judith Manz[6,7], Iain K. Pemberton[8], Jane MacDougall[8], Frances MK Williams[9], Jan Van Zundert[10], Klaas Buyse[10], Gordan Lauc[11], Ivan Gudelj[11], Dragan Primorac[12,13,14,15], Andrea Skelin[11,12], Yurii S. Aulchenko[17], Lennart C. Karssen[17], Leonardo Kapural[18], Richard Rauck[18], Guido Fanelli[1,2] on behalf of the PainOMICS Group*

PainOMICS Group* Collaborators

Melanie Waldenberger, PhD[6,7]

Manuela Zunino, MD[3]

Alice Montalti, MD[3]

Dario Bugada, MD[2]

Concetta Dagostino, PhD[2]

Silvana Montella, MD[2]

Marco Baciarello, MD[2]

Christian Compagnone, MD[2]

1 Department of surgical science. University of Parma, Parma, Italy.

2 Anesthesia Intensive Care and Pain Therapy service. Azienda Ospedaliera Universitaria Parma, Parma, Italy.

3 Anesthesia, Intensive Care and Pain Therapy, Emergency Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

4 Service of Biometry & Satistics, Research Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

5 School of Medical and Health Sciences, Edith Cowan University

6 Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany.

7 Institute of Epidemiology II, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany.

8 Photeomix, IP Research Consulting SAS, Noisy le Grand, France.

9 Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.

10 Department of Anesthesiology, Critical Care and Multidisciplinary Pain Center, Ziekenhuis Oost-Limburg, Genk, Belgium.

11 Genos Glycoscience Research Laboratory, Zagreb, Croatia

12 St. Catherine Specialty Hospital, Zabok, Croatia

13 Penn State University, Eberly College of Science, State College, PA, USA

14 Faculty of Medicine University of Osijek, Osijek, Croatia

15 University of Split School of Medicine, Split, Croatia

16 Children's Hospital Srebrnjak, Zagreb, Croatia

17 PolyOmica, Groningen, The Netherlands.

18 Carolinas Pain Institute, Winston-Salem, NC 27103, USA.

Corresponding Author:
Dr Massimo Allegri
Anesthesia Intensive Care and Pain Therapy service, Azienda Ospedaliera Universitaria Parma
via Gramsci 14, 43126 Parma (ITALY)
phone: 00390521703965
fax: 00390521703567
email: mallegri@parmanesthesia.com

Abstract

Introduction: Chronic low back pain (CLBP) produces considerable direct costs as well as indirect burdens for society, industry and health systems.

CLBP is characterized by heterogeneity, inclusion of several pain syndromes, different underlying molecular pathologies, and interaction with psychosocial factors that leads to a range of clinical manifestations. There is still much to understand in the underlying pathological processes and the non-psychosocial factors, which account for differences in outcomes.

Biomarkers that may be objectively used for diagnosis and personalized, targeted and cost-effective treatment are still lacking. Therefore, any data that may be obtained at the "-omics" level (glycomics, Activomics and genome-wide association studies - GWAS) may be helpful to use as dynamic biomarkers for elucidating CLBP pathogenesis and may ultimately provide prognostic information too.

By means of a retrospective, observational, case-cohort, multicenter study, we aim to investigate new promising biomarkers potentially able to solve some of the issues related to CLBP.

Methods and analysis: the study follows a two-phase, 1:2 case-control model. A total of 12000 individuals (4000 *cases* and 8000 *controls*) will be enrolled; clinical data will be registered, with particular attention to pain characteristics and outcomes of pain treatments. Blood samples will be collected to perform omics studies. The primary objective is to recognize genetic variants associated with CLBP; secondary objectives are to study glycomics and Activomics profiles associated with CLBP.

Ethics and dissemination: The study is part of the PainOMICS project funded by European Community in the 7th Framework Program. The study has been approved from competent Ethical Bodies and copies of approvals were provided to the European Commission before starting the study.

Results of the study will be reviewed by the Scientific Board and Ethical Committee of the PainOMICS Consortium.

The scientific results will be disseminated through peer-reviewed journals.

Registration details: Registered on Clinicaltrials.gov (NCT02037789).

Strengths and limitations of this study

Strength:

- Multiple-centre and multiple-discipline study: the study includes centres in the EU, USA and Australia, with research teams specialising in: (1) clinical aspects of pain, (2) biology and genetics of pain, (3) generation of 'omics' data, and (4) analysis of multiple 'omics' data.
- Hypothesis driven v.s. hypostasis generating: The study aims to profile "-omics" biomarkers (GWAS, glycomics and activomics) potentially to decipher the pathogenesis of CLBP associated with the different patho-physiological patterns.
- Longitudinal design with a large sample size: 1) Discovery phase 3000 cases and 6000 controls; 2) validation phase: 1000 cases and 2000 controls.

Limitations:

- While the heterogeneity of the study populations is helpful in the discovery phase, this may limit conclusions in the validation phase.
- Functional investigation with animal model has not been included in the current project due to the limitation of the funding.

Introduction

Low Back Pain (LBP) is one of the most common health problems worldwide with an estimated age-standardized point prevalence of 9.4%[1]. In 2012 a global review of the prevalence of low back pain was published reporting a mean ± SEM point prevalence of activity-limiting low back pain lasting more than 1 day of 11.9 ± 2.0%, and the 1-month prevalence of 23.2 ±2.9%[2]. LBP accounts for considerable disability and work absence, and ranks in the Global Burden of Disease 2013 study as a leading contributor to global disability measured in years lost due to disability (YLDs)[3]. LBP is defined as pain and discomfort, localized below the costal margin and above the inferior gluteal folds, with or without leg pain. Prognostic factors for LBP include demographic factors (educational attainment, age, gender), occupational factors (employment), mental health morbidity (anxiety, depression), perception of pain and disability (pain intensity and expectation of persistent pain) and other psychological factors (fear avoidance, catastrophizing, illness perceptions) [4].

LBP becomes chronic low back pain (CLBP) when symptoms last at least 3 months. Activity-limiting LBP tends to recur and the course of LBP is increasingly viewed as a chronic recurring condition[5], which accounts for considerable direct economic costs as well as indirect burdens for society, industry and health systems[6, 7]. The prevalence of chronic low back pain (CLBP) appears to be rising with an increase from 3.9% in 1992 to 10.2% in 2006 in the USA [8], and from 4.2% in 2002 to 9.6% in 2010 in Brazil[9].

Pain is a subjective sensation, which is influenced by a range of physical and psychosocial factors through poorly understood neural mechanisms[10]. The advances in medical imaging have improved our ability to identify the anatomic origin of CLBP, however CLBP is heterogeneous and the anatomic site only tells part of the story. There is considerable variation in prognosis and, at present, CLBP exerts a substantial burden on the individual, the family and workplace.

Evidence based guidelines recommend the initial exclusion of serious diagnosis before the implementation of clinical management, which promotes continued function and best practice rehabilitation approaches. However, there is much more to understand about the underlying process, how this affects the prognosis and how we can use this to tailor treatment for the individual. In making a diagnosis of LBP, red flag symptoms are used to identify the need for investigation for underlying serious illness[11]. Yellow flags (psychosocial factors) identify risks of chronicity [12]and highlight the heterogeneity and complexity of CLBP where the severity, chronicity and prognosis may depend on the anatomical site, the underlying pathological process, comorbidities as well as individual psychosocial factors.

While psychosocial factors clearly influence outcomes in CLBP, genetic and epigenetic factors may account for some variation in response to treatment. Even though both persistent CLBP and disc degeneration are known to be heritable[13, 14] and the two traits are highly related to one another, with disc degeneration being a major predictor for LBP episodes [14], few genetic variants have been identified and confirmed for both traits [15-17]. Only two genome wide association studies report on chronic/persistent widespread pain [16, 17] and two GWAS of intervertebral disc degeneration [18, 19]. In keeping with other common complex traits, the individual effects of the identified loci are small and explain only a small fraction of the trait or disease variation [20]. As such, they do not substantially improve predictions over those based on known factors such as family history [21-23]. Unfortunately these data have not yet shed light on the pain pathogenesis mechanisms and they do not offer prediction of treatment likely suitable in individual patients. Replication of these findings is also needed. Hence, despite promising recent data, new studies are needed to identify objective biomarkers for both diagnosis and prediction of treatment's efficacy in CLBP patients.

Glycomics is an emerging field, recently identified as a priority for the next decade by the US National Academies of Science [24, 25]. Recent studies reported on protein glycosylation in large human population samples, with promising glycan profiling for disease diagnosis and stratification e.g., autoimmune diseases and hematological cancers, metabolic syndrome, systemic lupus erythematosus and many other diseases [26-33]

Together with glycomics Activomics may be also useful in detecting new biomarkers for low back pain as it combines data involved in enzymatic activity of several post-translational modification proteins in an integrated model, providing dynamic characterization of the current state of an organism.

Activomics®, is a novel –omic strategy that aims to describe biological systems in terms of differential protein post-translational modification (PTM) activities. Perturbations in intracellular and/or intercellular cell signalling networks are frequently linked to chronic diseases such as cancer. Enzyme activities in serum are monitored using a proprietary panel of protein and peptide substrates under multiple assay conditions

including the judicious use of enzyme cofactors and inhibitors in order to optimize the discrimination between protein modification enzymes with preferences for overlapping primary sequence or structural targets. Principal Activomics substrate panels include proteases (metalloproteinases, serine, cysteine, aspartic proteases) and their protease inhibitors (e.g. serpins), caspases, kinases (ser/thr and tyr), phosphatases and (de)acetylases.

Here we present the study protocol for a retrospective analysis, in a large cohort of CLBP patients, to determine "-omics" biomarkers (GWAS, glycomics and Activomics) potentially associated with susceptibility to CLBP and with different patho-physiological patterns.
We will link clinical data to the multiple "-omics" analyses, thereby profiling novel biomarkers, which may advance our knowledge of some of the remaining unsolved problems in CLBP.

Methods

We present a retrospective observational, multicenter, international clinical study, with a case control design.
The study is part of the PainOMICS project that includes four different trials and was reviewed and funded by European Community in the 7th Framework Program (FP7) - THEME [HEALTH.2013.2.2.1-5 - Understanding and controlling pain].
The project includes six clinical centers from Italy, Croatia, Belgium, Australia and the United States and four centers for scientific analyses in Croatia, France, Germany and the United Kingdom. Statistical expertise is provided by the "PolyOmica" consulting based in the Netherlands.
The retrospective study was approved by the ethical committees of each separate clinical center between December 2013 and March 2014 and patient enrolment is currently active up to September 2016.
The study is registered at Clinicaltrials.gov (NCT02037789).

Participant enrolment and data collection

Figure 1 Study Flow Chart

Cases (patients having CLBP) will be enrolled by each participating center. Every effort will be made to accumulate a well characterized cohort of patients with persistent CLBP, sub-grouped according to the likely anatomical cause of the pain. Patients fulfilling the following conditions will be considered for enrolment: age 18 years and older; chronic pain (pain lasting longer than 12 weeks) between the costal margins and gluteal fold, with or without symptoms in one or both legs, written informed consent signed and Caucasian ancestry.
Controls (patients without chronic low back pain) will be retrieved from two different sources: 1) existing bio-banks of healthy subjects having collected information about CLBP; 2) subjects enrolled in the parallel prospective study on acute LBP (part of PainOMICS project – NCT 02037763), i.e. patients who presented

with acute LBP and have not become chronic over 6 months but the pain has resolved. Age (decades) and gender distribution of controls will be matched as closely as possible to that of the cases. Controls are enrolled according to the following inclusion criteria: older than 18 years, no chronic pain (lasting longer than 12 weeks) in the past 12 months, written informed consent obtained and Caucasian ancestry.

Subjects with any evidence of clinically unstable disease, severe psychiatric disorder (excluding mild depression) or mental impairment; recent history (less than one year) of spinal fracture, back pain due to spinal tumor or infection, pregnancy, will be excluded from the study.

Patients and controls selected for participation will receive a detailed description of the study and will be asked to sign an informed consent prior to entering the study (Enrolment Visit).

Once enrolled in the trial, patients will be assigned to a unique anonymous code. Data collection includes; demographics (age, gender, race, body mass index - BMI, occupational history), clinical and pharmacological history, pain characteristics (onset, duration, intensity, pain referral pattern, irradiation, sensory abnormalities, precipitating events, history of previous episodes), effectiveness/tolerability of pain treatments received (when applicable). A specific questionnaire (Pain-DETECT, PD) is applied to evaluate both pain type and the pain generator, as well as the possible pathophysiological (nociceptive and/or neuropathic) mechanism sustaining CLBP and functional impairment.

Blood samples for omics analyses are taken from each enrolled patient at the time of the enrolment consultation, and biological samples are sent to analytical partners of the Consortium for specific omics analysis.

Clinical data are collected in the designated ad-hoc Case Report Form and into a dedicated web database (REDCap - Research Electronic Data Capture); access to the web-database is restricted to the project partners and can be accessed using a dedicated username and a password.

Omics data are centralized in a specific supervised database.

Samples collection methodology

All the clinical centers have to guarantee that the biological samples from each patient will be prepared, stored, and shipped following the analytical procedures described in three standard operating procedures (SOPs) developed and validated to provide details for conducting the study, phase by phase, with written instructions to achieve uniformity of the procedures used for obtaining patient blood for omic analysis techniques, storing the samples, and shipping the aliquots to the specialized laboratories.

All the patients and controls undergo blood sampling for omics determinations. The samples will then be divided into two tubes containing EDTA for genetic and glycomic analyses and into one serum tube with clot activator plus gel for the Activomics.

Detailed description of the validation of SOPs is provided in a separate paper under submission.

Analyses for biologic markers

Genetic analyses:

GWAS analyses will be performed using genome-wide Illumina genotyping technology. For each genetic variant a standard association model (like linear or logistic regression model) will be used to investigate associations between molecular phenotypes and genetic markers. Stringent Bonferroni correction will be applied to determine a genome-wide significance level at a nominal alpha-level of 0.05. The results of the separate GWAS will be meta-analyzed applying a standard model such as inverse variance weighting. In addition independent single-nucleotide polymorphisms (SNPs) will be investigated for in the same region as primary SNPs. Additional associated SNPs will be identified by implementation of functional association networks constructed from databases, mostly containing protein-protein interaction information.

To allow evaluating the procedure, the explained variances of separate SNPs and the complete set of markers will be assessed.

Glycomics analyses:

Glycomics analyses will be performed on total serum proteins and on a single serum protein, immunoglobulin G (IgG). IgG will be isolated from serum samples by affinity chromatography using 96-well monolithic plates with Protein G as previously described (Pucic et al, MCP, 2011). N-glycans will be released from total serum proteins and IgG by overnight deglycosylation with N-glycosidase F (PNGase F). Released N-glycans will be fluorescently labeled with 2-aminobenzamide (2-AB) fluorescent tag and purified by hydrophilic interaction liquid chromatography (HILIC) solid phase extraction (SPE). Labeled N-glycans will be analyzed by hydrophilic interaction chromatography on a Waters Acquity UPLC instrument using Waters BEH Glycan chromatography column, 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. The system will be calibrated using an external standard of hydrolyzed and 2-AB labeled glucose oligomers from which the retention times for the individual glycans will be converted to glucose units. Data processing will be performed with an automatic processing method after which each chromatogram will be manually corrected to maintain the same intervals of integration for all the samples. The chromatograms obtained will all be separated in the same manner into peaks and the amount of glycans in each peak will be expressed as percentage of total integrated area.

Activomics analyses:

Activomics analyses will be performed on retrospective serum samples from anonymous but well characterized patients. Samples will be collected, handled and analyzed in a way that minimizes freeze-thaw cycles. For each enzymatic reaction tested, 1 – 20 microliters of serum from patients and healthy controls will be incubated with the appropriate Activomics® substrate under controlled conditions (time, temperature, optimized buffer conditions, etc.). Enzymatic modification of the substrate will be monitored quantitatively by proprietary charge-based microfluidic assays for subsequent statistical analysis. The assays will be repeated for a panel of different substrates in order to provide a wide view of disease-related changes to post-translational modification activities for multivariate statistical analysis. Performance characteristics of the

panel will be assessed by univariate and multivariate hierarchical clustering and principal component analysis to differentiate activities of disease versus control samples. Sensitivity and specificity will be evaluated in ROC analyses to define cut-off values in the design of the optimal predictive biomarker panel.

Primary and secondary objectives

The primary objective of this retrospective study is to recognize genetic variants associated with persistent CLBP, by comparing CLBP patients and pain-free patients. We will correlate genetic variants associated with CLBP through a GWAS study, in a wide international population of European ancestry.

Secondary objectives are to recognize glycomic and Activomic data associated with CLBP patients compared to patients without CLBP.

The participating clinical centers have defined a minimal shared diagnostic dataset available in all clinical centers. Each of the participants will stratify patients according to their clinical features, imaging data and results from Pain-DETECT. Considering the patient's response to diagnostic procedures, patients will be sub-grouped (taking into account the patient history, clinical examination, radiological results and potentially the response to diagnostic blocks) into 6 main categories:

1. spinal stenosis,
2. discogenic pain,
3. facet joint pain,
4. sacroiliac joint pain,
5. low back pain with radicular pain (not predominant radicular pain),
6. widespread low back pain.

Statistics

Study design

The study will follow a two-phase (discovery & validation), 1:2 case-control model:

- Discovery population: random sample of two thirds of the entire population of cases.
- Validation population: the remaining one third. Following the GWAS phase, the genes discovered will be assessed for biological plausibility and entered into the validation phase

Sample size calculation

Since, to date, no data on the omics of CLBP are available in literature, and since the variants associated with most common diseases have modest effects we considered a number of scenarios, ranging in model assumptions with respect to allele frequency and effect.

With 3000 cases and 6000 controls, we assessed genetic scenarios in which we have 80% power to detect association at genome-wide significant level[34]. Consistent with the literature, for high odds ratios (OR=2) we will be able to detect variants with low minor allele frequency (MAF≥1.5%). With higher allele

frequency we will be able to find smaller effects; for example, for variants with MAF=25% we have power to detect OR≥1.25

For the replication/validation phase, with 1000 cases and 2000 controls, we have 80% power to confirm detected variants when using nominal p = 0.005 (leading to experiment-wise type I error of 0.05 assuming 10 tests) and approximately 60% power in case of more severe multiple testing (100 tests, leading to nominal p = 0.0005).

Statistical analysis

All the enrolled patients and controls will be analyzed.

Discovery phase: For analyses investigating highly dimensional omics space we will use a range of approaches. The first approach will extend the classical sequential framework. Predictor screening will be performed by logistic or Cox regression models traversing through the omics space and incorporating few predictors at a time. Statistically significant (at experiment-wise level) predictors will be included in the model, and the next iteration through the omics space will be performed. A classical example of this approach includes genome-wide association analysis, followed by conditional analyses for identification of secondary signals. To investigate a large numbers of predictors simultaneously, we will use modern regularization/shrinkage and machine learning methods allowing analysis of (relatively) large numbers of predictors jointly. While this type of approach does not address the question of statistical testing in the same way as "classical" approaches do, it is widely used in the context of biomarker discovery, where prediction and not the *p*-values are of primary interest. For all methods aimed towards biomarker discovery, the accuracy of prediction will be accessed by cross-validation, and optimal solutions will be analyzed to identify potential biomarkers, which will be selected on their discriminative value in a receiver operator characteristic (ROC) analysis.

Validation phase: The discovered genetic variants will be examined bio-informatically for biological plausibility before entering the validation phase. The association of the candidate polymorphisms with the outcome (being a case) will be assessed with logistic regression. The following strategies will be used: single genes assessment/genetic score (sum of candidate genes)/multiple genes. Adjustment for covariates (age, gender, clinical features) will be performed.

Details of the statistical analyses will be provided in the final statistical analysis plan (SAP).

The analysis of the secondary endpoints will follow the same principles reported above.

Ethical Issues

Sample collection and use of clinical data has started only after the Ethical Approval of the present study protocol from the competent ethical bodies (Ethics Committees of the Institutions involved in patients enrolment).

Copies of Ethical Approval were provided to the European Commission before initiating the study. All protocol, copies of Informed Consent and Information Sheets approved by the competent ethical bodies, were provided to the European Commission before starting the study.

The Scientific Board and Ethical Committee of PainOMICS Consortium will also review the results of the study in order to evaluate any possible societal impact of our findings according to the ethical concerns about genetics/OMICS and diagnosis of chronic pain [35].

Monitoring and quality assessment

Patients will be withdrawn from the study in case of withdrawal of consent (subjects may always and without obligation withdraw their informed consent), or any other condition that, upon clinical judgment of the investigator, will make unacceptable further study participation for that individual patient.

The Coordinating Investigator (University Hospital of Parma, Italy) will delegate, in each participating center, a clinical supervisor (to ensure that the study is conducted according to the protocol, to good clinical practice, and to national regulations) and also a data monitor, to ensure accuracy, completeness and verification of patients' data. The data monitors, from each participating center, will make up the data monitoring committee. The External Project Advisory Committee (EPAC) of the Pain-OMICS FP7 project will perform an overall scientific supervision of the trial and of the emergent data.

The participating members will discuss results and any issues of the study at regular audits during the annual SIMPAR (Study in Multidisciplinary Pain Research) meeting, and in any other case that may be deemed necessary.

Conclusions

This study is, to the best of our knowledge, the first to investigate genetic and OMIC biomarkers in a large population sample of CLPB patients. These biomarkers may be related to pain sensation, as well as to disease pathophysiology and pain generators. The overall objective is to validate associations that may result in a more personalized diagnosis and therapy of a disease with a high health and societal burden such LBP.

Furthermore, the novel biomarkers emerging from this retrospective study will be validated in a prospective cohort collected within the same PainOMICS project, in order to evaluate their ability to predict the possibility of advancement to chronic pain in patients suffering from an acute episode of LBP.

The Pain-OMICS project is expected to significantly expand the level of knowledge on how low back pain is generated, propagated and quenched. We will mobilize significant human and material resources in Europe and USA, allowing a comprehensive characterization of large cohorts of CLPB patients, aiming to identify a number of potential biomarkers related to different aspects of chronic low back pain, as well as potential new targets for therapy.

Authors' contributions:

MA, CG, IKP, FMKW, JVZ, GL, DP, YSA, LK and GF conceived the study and revised the paper.

MA, CEM drafted the paper. MDG, CK, WW, MS, JM, JMD, KB, IG, AS, LCK, RR provided feedback on the manuscript, and all authors reviewed and approved the final version of the paper.

Acknowledgements

MW and CD provided scientific support and are involved in sample management and analysis.

MZ, AM, DB, SM, MB, CC are collecting data and caring for study patients.

Funding statement

This trial is funded academic/SME research; it is supported by funding from the European Commission in the context of the Seventh Framework Program of the European Community for Research, Technological Development and Demonstration Activities - (FP7) - THEME [HEALTH.2013.2.2.1-5 - Understanding and controlling pain].

The present study is not funded by Industry or any other commercial sponsors.

Competing Interests

The study is supported by a grant from the European Commission (602736).

Dr Allegri is a consultant for Grunenthal, Angelini and Mundipharma. He also collaborated for speeches with MSD and Carefusion.

Dr. Lauc has multiple patents in the field of glycoscience issued.

Dr. Aulchenko is a director and co-owner of Maatschap PolyOmica, which provides (consulting) services in the area of (statistical) (gen)omics.

Data sharing statement

No additional data are available

References

1. Hoy, D., et al., *The global burden of low back pain: estimates from the Global Burden of Disease 2010 study.* Ann Rheum Dis, 2014. **73**(6): p. 968-74.
2. Hoy, D., et al., *A systematic review of the global prevalence of low back pain.* Arthritis Rheum, 2012. **64**(6): p. 2028-37.
3. Collaborators, G.B.o.D.S., *Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.* Lancet, 2015. **386**(9995): p. 743-800.
4. Campbell, P., et al., *Prognostic indicators of low back pain in primary care: five-year prospective study.* J Pain, 2013. **14**(8): p. 873-83.

5. Hoy, D., et al., *The Epidemiology of low back pain.* Best Pract Res Clin Rheumatol, 2010. **24**(6): p. 769-81.
6. Steenstra, I.A., et al., *Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature.* Occup Environ Med, 2005. **62**(12): p. 851-60.
7. Organization, W.H., *Priority medicines for Europe and the worldupdate report 2013. Chapter 6.*
8. Freburger, J.K., et al., *The rising prevalence of chronic low back pain.* Arch Intern Med, 2009. **169**(3): p. 251-8.
9. Meucci, R.D., et al., *Increase of chronic low back pain prevalence in a medium-sized city of southern Brazil.* BMC Musculoskelet Disord, 2013. **14**: p. 155.
10. Koyama, T., et al., *The subjective experience of pain: where expectations become reality.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12950-5.
11. http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/red-flags-in-back-pain. Available from: http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/red-flags-in-back-pain.
12. http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/yellow-flags-in-back-pain. Available from: 12. http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/yellow-flags-in-back-pain.
13. Battie, M.C., et al., *Heritability of low back pain and the role of disc degeneration.* Pain, 2007. **131**(3): p. 272-80.
14. MacGregor, A.J., et al., *Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins.* Arthritis Rheum, 2004. **51**(2): p. 160-7.
15. Eskola, P.J., et al., *Genetic association studies in lumbar disc degeneration: a systematic review.* PLoS One, 2012. **7**(11): p. e49995.
16. Kim, H., D. Clark, and R.A. Dionne, *Genetic contributions to clinical pain and analgesia: avoiding pitfalls in genetic research.* J Pain, 2009. **10**(7): p. 663-93.
17. Peters, M.J., et al., *Genome-wide association study meta-analysis of chronic widespread pain: evidence for involvement of the 5p15.2 region.* Ann Rheum Dis, 2013. **72**(3): p. 427-36.
18. Williams, F.M., et al., *Novel genetic variants associated with lumbar disc degeneration in northern Europeans: a meta-analysis of 4600 subjects.* Ann Rheum Dis, 2013. **72**(7): p. 1141-8.
19. Song, Y.Q., et al., *Lumbar disc degeneration is linked to a carbohydrate sulfotransferase 3 variant.* J Clin Invest, 2013. **123**(11): p. 4909-17.
20. Visscher, P.M., et al., *Five years of GWAS discovery.* Am J Hum Genet, 2012. **90**(1): p. 7-24.
21. Aulchenko, Y.S., et al., *Predicting human height by Victorian and genomic methods.* Eur J Hum Genet, 2009. **17**(8): p. 1070-5.
22. Clayton, D.G., *Prediction and interaction in complex disease genetics: experience in type 1 diabetes.* PLoS Genet, 2009. **5**(7): p. e1000540.
23. Eichler, E.E., et al., *Missing heritability and strategies for finding the underlying causes of complex disease.* Nat Rev Genet, 2010. **11**(6): p. 446-50.
24. *National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. Transforming Glycoscience: A Roadmap for the Future.* 2012; available on http://www.ncbi.nlm.nih.gov/books/NBK109958/, Washington (DC): National Academies Press (US).
25. Lauc, G., et al., *Complex genetic regulation of protein glycosylation.* Mol Biosyst, 2010. **6**(2): p. 329-35.
26. Knezevic, A., et al., *Variability, heritability and environmental determinants of human plasma N-glycome.* J Proteome Res, 2009. **8**(2): p. 694-701.

27. Lauc, G., et al., *Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers.* PLoS Genet, 2013. **9**(1): p. e1003225.
28. Lauc, G., et al., *Mechanisms of disease: The human N-glycome.* Biochim Biophys Acta, 2015.
29. Lu, J.P., et al., *Screening novel biomarkers for metabolic syndrome by profiling human plasma N-glycans in Chinese Han and Croatian populations.* J Proteome Res, 2011. **10**(11): p. 4959-69.
30. Pabst, M., et al., *IL-1beta and TNF-alpha alter the glycophenotype of primary human chondrocytes in vitro.* Carbohydr Res, 2010. **345**(10): p. 1389-93.
31. Pucic, M., et al., *High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations.* Mol Cell Proteomics, 2011. **10**(10): p. M111 010090.
32. Ruhaak, L.R., et al., *Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity.* PLoS One, 2010. **5**(9): p. e12566.
33. Vuckovic, F., et al., *Systemic lupus erythematosus associates with the decreased immunosuppressive potential of the IgG glycome.* Arthritis Rheumatol, 2015.
34. Pe'er, I., et al., *Estimation of the multiple testing burden for genomewide association studies of nearly all common variants.* Genet Epidemiol, 2008. **32**(4): p. 381-5.
35. Compagnone, C., et al., *Ethical issues in pain and omics research. Some points to start the debate.* Croat Med J, 2014. **55**(1): p. 1-2.

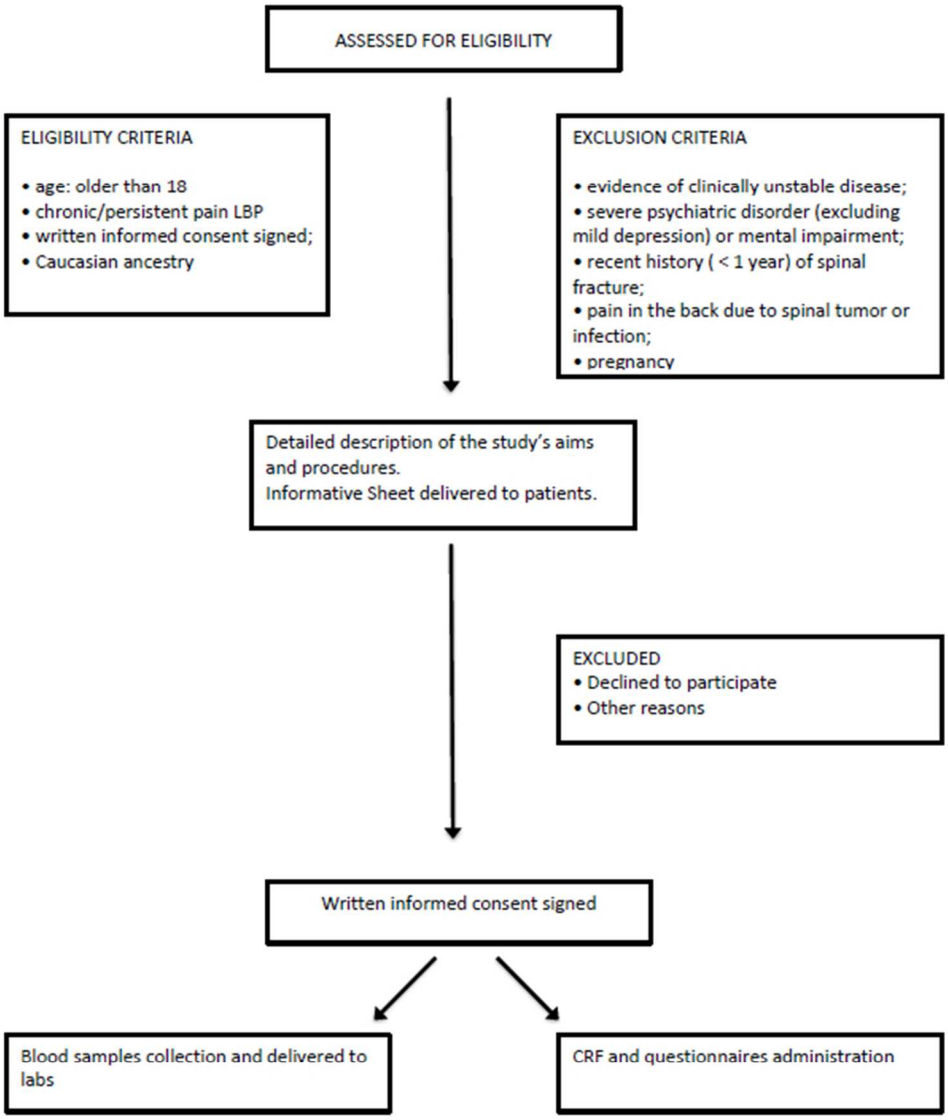

Figure 1 Study Flow Chart

163x199mm (96 x 96 DPI)

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description
Administrative information		
Title	1	"OMICS" BIOMARKERS ASSOCIATED WITH CHRONIC LOW BACK PAIN: A RETROSPECTIVE CLINICAL STUDY
Trial registration	2a	Clinicaltrials.gov (NCT02037789).
	2b	All items from the World Health Organization Trial Registration Data Set: **Listed below the present checklist**
Protocol version	3	Version I, 10/23/2013: original version approved Version II 01/24/2014: amendment 01 for minor corrections in the text Version III 02/18/2015: amendment 02 following the transfer of Coordinator (Dr Massimo Allegri) and Coordinating activity from Fondazione IRCCS Policlinico San Matteo, Pavia, Italy to Azienda Ospedaliero-Universitaria di Parma, Italy. Each amendment has been approved by Ethical Committees at the interested clinical centers.
Funding	4	The present study is not funded by Industry or any other commercial sponsors. This trial is a funded academic research; it is supported by founding from the European Commission in the context of the Seventh Framework Program of the European Community for Research, Technological Development and Demonstration Activities. Grant agreement no: 602736.

Roles and responsibilities	5a	**Coordinating Investigator:** Massimo Allegri,MD. Department of Surgical Science, University of Parma, Parma, Italy **Clinical Sites of recruitment and local principal investigators:** Dr. Cristina E. Minella: Fondazione IRCCS Policlinico San Matteo, Pavia, Italy Dr Jan Van Zundert: Multidisciplinary Pain Centre, Hospital Oost--Limburg (ZOL), Genk , Belgium Professor Dragan Primorac: "St.Catharine" Orthopedics, Surgery, Neurology and Physical Medicine and Rehabilitation Specialty Hospital (St-Cat), Zabok, Croatia Dr Leonardo Kapural: The Center for Clinical Research (CPI), Winston-Salem, USA Professor Wei Wang: Edith Cowan University (ECU), Perth, Australia Participating laboratories Professor Gordan Lauc: Genos Glycoscience Research Laboratory, Zagreb, Croatia Professor Christian Gieger: Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany. Dr Iain K Pemberton: Photeomix, IP Research Consulting SAS, Noisy le Grand, France Dr Frances MK Williams: Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. **Bioinformatics center:** Professor Yurii S Aulchenko: PolyOmica, Groningen, The Netherlands. MA, CEM, JVZ, DP, LK, WW, GL, CG, IKP, FMKW and GF conceived the study and implemented the protocol, YSA provided statistical expertise in clinical trial design.
	5b	Name and contact information for the trial sponsor: **N/A**
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities: **N/A**

5d	The Coordinating Investigator is the responsible for validity, completeness, exactness and plausibility of data recorded in the CRFs, and for their correspondence to data recorded in subject's medical records. The Coordinating Investigator delegates, in each participant centre, a clinical supervisor (in charge of ensuring that the study will be conducted according to the protocol, to the good clinical practices, and to national regulations) and a data monitor, to ensure accuracy, completeness and verification of patient's data. Data monitoring committee: all data monitors, delegated at participating centres, will constitute the data monitoring committee. Scientific and Steering committee: the External Project Advisory Committee (EPAC) of pain-OMICS FP7project will perform an overall scientific supervision of the trial and of emerging data.

Introduction

Background and rationale	6a	Low back pain (LBP) is one of the most common medical problems encountered in daily life; it is related to disability and work absence and accounts for high economical costs in Western societies. Low-back pain is a diverse group of mixed pain syndromes (neuropathic and nociceptive) with different molecular pathologies at different structural levels displaying similar clinical manifestations. Currently, there are limited biomarkers (mostly imaging) or clinical findings that can be used objectively to help the physician in precise anatomic diagnosis leading to the safest and most cost-effective treatment for the patient (reduction of direct and indirect costs and improvement of treatment efficacy). The main aim of this trial is to identify all "omics biomarkers" associated with susceptibility to chronic/persistent LBP and its different pathophysiology. Investigators will compare "omic biomarkers" between patients with and without persistent chronic low back pain (CLBP). "OMIC" biomarkers investigated will be genetics, glycomics and activomic. Genetics through GWA studies has already obtained important results in pain research; however concerning low back pain, there is not yet suitable genotype-phenotype correlations helpful to stratify patients. Glycomics is an emerging field that has recently been identified as a priority for the next decade by the US National Academies of Science. Many common complex diseases will be associated with specific changes in glycan structures. In addition, common genetic polymorphisms influencing glycosylation and consequent differences in glycome composition could be important diagnostic and prognostic markers. The first studies reporting protein glycosylation in large human population samples have been recently published by partners in the consortium. Reliable identification of valid associations between specific glyco-phenotypes and predisposition for the development or progression of a specific disease requires analysis of thousands of patients. Activomics: combines data about enzymatic activity of numerous numerous post-translational modification proteins in an integrated model which provides dynamic characterization of the current state of an organism. In this project information about numerous proteases, kinases, phosphatases and glycosidases will be collected and used to complement the existing phenotype information.
	6b	Explanation for choice of comparators

Objectives	7	The primary objective is to recognize genetic variants associated with persistent CLBP patients compared to patients without chronic/persistent pain. Through a Genetic Wide Association Study (GWAS) we will correlate genetic variants associated with persistent CLBP in a wide, international population of European ancestry. Secondary objectives 1. Recognize Glycomic and Activomic data associated with persistent CLBP patients compared to patients without chronic/persistent pain. Glycomic: Glycome composition is very variable between individuals and different glycosylation can significantly affect function of various proteins. Plasma glycome is expected to be relevant for generation and transmission of pain either as direct reflection of the immune cell function, or as a proxy for glycosylation of membrane proteins in different signalling pathways. By analysing composition of the plasma glycome in a large number of patients with low back pain, we will identify glyco-phenotypes which may be associated with individual variation in the way pain is generated, transmitted, or quenched. Activomic: the objective will be to define a panel of putative biomarkers based on protein post-translational modification enzymatic activities present in patient samples that can be differentiated statistically from control samples in order to diagnose and stratify lower back pain and/or its different phenotypes 2. All omic data will be compared stratifying our population according to: - Pathophysiology: discogenic pain, spinal stenosis, facet joint pain, sacroiliac joint pain, low back pain with radicular pain (not predominant radicular pain) and widespread pain. - pain intensity - response to treatment - duration of pain
Trial design	8	Retrospective observational multinational clinical study, with a case control design.

Methods: Participants, interventions, and outcomes

Study setting	9	Study settings: community clinics and academic hospitals in the clinical participating centres listed above (Italy, Croatia, Belgium, USA and Australia). Cases are collected at each participating centre. Every effort will be made to accumulate a well phenotyped cohort of patients with persistent CLBP, sub-grouped into 6 categories: discogenic pain, spinal stenosis (congenital or acquired), facet joint pain, sacroiliac joint pain, low back pain with radicular pain (not predominant radicular pain) and widespread low back pain. Controls (patients without chronic low back pain) will be retrieved from two different sources: 1) existing bio-banks of healthy subjects having collected information about CLBP; 2) subjects enrolled in the parallel prospective study on acute LBP (part of PainOMICS project – NCT 02037763), i.e. patients who presented with acute LBP and have not become chronic over 6 months but the pain has resolved. The participating clinical centres will identify minimal diagnostic dataset available in all six clinical centres that will be sufficient to stratify persistent CLBP patients according to the origin of pain, progression and the response to therapy. This harmonization of clinical definition and stratification of patients will create a framework for the correlation of well phenotyped subjects with “omics” results.
Eligibility criteria	10	Inclusion Criteria of patients with persistent CLBP: age: older than 18; chronic/persistent pain (pain lasting longer than 12 weeks) between the costal margins and gluteal fold, with or without symptoms into one or both legs; written informed consent signed; Caucasian ancestry Inclusion Criteria of healthy volunteers: age: older than 18; without any chronic/persistent pain (pain lasting longer than 12 weeks) in the last one year; written informed consent signed; Caucasian ancestry Exclusion Criteria: evidence of clinically unstable disease; severe psychiatric disorder (excluding mild depression) or mental impairment; recent history (< 1 year) of spinal fracture; pain in the back due to spinal tumor or infection; pregnancy.

Interventions	11a	Patients and controls selected for participation will receive a detailed description of the study and will be asked to sign an informed consent prior to entering the study (Enrolment Visit). Once enrolled in the trial, individuals will be assigned to a unique anonymous code. Data collection includes; demographics (age, gender, race, body mass index - BMI, occupational history), clinical and pharmacological history, pain characteristics (onset, duration, intensity, pain referral pattern, irradiation, sensory abnormalities, precipitating events, history of previous episodes), effectiveness/tolerability of pain treatments received (when applicable). A specific questionnaire (Pain-DETECT, PD) is applied. Blood samples for omics analyses are taken from each enrolled patient at the time of the enrolment consultation, and biological samples are sent to analytical partners of the Consortium for specific omics analysis. Clinical data are collected in the designated ad-hoc Case Report Form and into a dedicated web database (REDCap - Research Electronic Data Capture). Omics data are centralized in a specific supervised database.
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease) **N/A**
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests) **N/A**
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial **N/A**

Outcomes	12	The primary objective of this retrospective study is to recognize genetic variants associated with persistent CLBP, by comparing CLBP patients and pain-free patients. We will correlate genetic variants associated with CLBP through a GWAS study, in a wide international population of European ancestry. Secondary objectives are to recognize glycomic and Activomic data associated with CLBP patients compared to patients without CLBP. The participating clinical centers have defined a minimal shared diagnostic dataset available in all clinical centers. Each of the participants will stratify patients according to their clinical features, imaging data and results from Pain-DETECT. Considering the patient's response to diagnostic procedures, patients will be sub-grouped (taking into account the patient history, clinical examination, radiological results and potentially the response to diagnostic blocks) into 6 main categories: - spinal stenosis, - discogenic pain, - facet joint pain, - sacroiliac joint pain, - low back pain with radicular pain (not predominant radicular pain), - widespread low back pain.
Participant timeline	13	Sample collection and use of clinical data started only after the Ethical Approval of the present study protocol from the competent ethical bodies (Ethics Committees of the Institutions involved in patients enrolment), and after the Administrative Approval. Copies of Ethical Approval have been provided to the European Commission before initiating the study. All protocol, copies of Informed Consent and Information Sheets once approved by the competent ethical bodies, have been provided to the European Commission before initiating the study. Enrolment is currently active up to September 2016.
Sample size	14	Since, to date, no data on the omics of CLBP are available in literature, and since the variants associated with most common diseases have modest effects we considered a number of scenarios, ranging in model assumptions with respect to allele frequency and effect. With 3000 cases and 6000 controls, we assessed genetic scenarios in which we have 80% power to detect association at genome-wide significant level. Consistent with the literature, for high odds ratios (OR=2) we will be able to detect variants with low minor allele frequency (MAF>1.5%). With higher allele frequency we will be able to find smaller effects; for example, for variants with MAF=25% we have power to detect OR>1.25 For the replication/validation phase, with 1000 cases and 2000 controls, we have 80% power to confirm detected variants when using nominal p = 0.005 (leading to experiment-wise type I error of 0.05 assuming 10 tests) and approximately 60% power in case of more severe multiple testing (100 tests, leading to nominal p = 0.0005).

Recruitment	15	The participation of six large clinical centres of the PAIN-OMICS Consortium will enable the identification of novel individual and composite biomarkers. The participating clinical Centres are leaders in treating patients with low-back pain, and are treating over 4,000 new patients with persistent CLBP each year, thus a large cohort of patients will be available for the collection of clinical data and biological samples.

Methods: Assignment of interventions (for controlled trials)

Allocation:		
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions **N/A**
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned **N/A**
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions **N/A**
Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how **N/A**
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial **N/A**

Methods: Data collection, management, and analysis

Data collection methods	18a	Patients will be enrolled during a medical consultation for CLBP, and after providing informed consent: clinical and demographic data will be registered, with particular attention to pain characteristics and effectiveness/tolerability of pain treatments received; a specific questionnaire (Pain-DETECT) will be applied. Blood samples will be collected from each patient and control enrolled in the study to perform glycomics, Activomics and GWAS studies. All the clinical centers have to guarantee that the biological samples from each patient will be prepared, stored, and shipped following the analytical procedures described in three standard operating procedures (SOPs) developed and validated to provide details for conducting the study, phase by phase, with written instructions to achieve uniformity of the procedures used for obtaining patient blood for omic analysis techniques, storing the samples, and shipping the aliquots to the specialized laboratories. All the patients and controls undergo blood sampling for omics determinations. The samples will then be divided into two tubes containing EDTA for genetic and glycomic analyses and into one serum tube with clot activator plus gel for the Activomics.
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols **N/A**
Data management	19	Once enrolled in the trial, individuals are assigned to a unique anonymous code. Clinical data are collected in the designated ad-hoc Case Report Form and into a dedicated web database (REDCap - Research Electronic Data Capture); access to the web-database is restricted to the project partners and can be accessed using a dedicated username and a password. Omics data are centralized in a specific supervised database.

Statistical methods	20a	All the enrolled patients and controls will be analyzed. Discovery phase: For analyses investigating highly dimensional omics space we will use a range of approaches. The first approach will extend the classical sequential framework. Predictor screening will be performed by logistic or Cox regression models traversing through the omics space and incorporating few predictors at a time. Statistically significant (at experiment-wise level) predictors will be included in the model, and the next iteration through the omics space will be performed. A classical example of this approach includes genome-wide association analysis, followed by conditional analyses for identification of secondary signals. To investigate a large numbers of predictors simultaneously, we will use modern regularization/shrinkage and machine learning methods allowing analysis of (relatively) large numbers of predictors jointly. While this type of approach does not address the question of statistical testing in the same way as “classical” approaches do, it is widely used in the context of biomarker discovery, where prediction and not the p-values are of primary interest. For all methods aimed towards biomarker discovery, the accuracy of prediction will be accessed by cross-validation, and optimal solutions will be analyzed to identify potential biomarkers, which will be selected on their discriminative value in a receiver operator characteristic (ROC) analysis. Validation phase: The discovered genetic variants will be examined bio-informatically for biological plausibility before entering the validation phase. The association of the candidate polymorphisms with the outcome (being a case) will be assessed with logistic regression. The following strategies will be used: single genes assessment/genetic score (sum of candidate genes)/multiple genes. Adjustment for covariates (age, gender, clinical features) will be performed. Details of the statistical analyses will be provided in the final statistical analysis plan (SAP). The analysis of the secondary endpoints will follow the same principles reported above.
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)

Methods: Monitoring

Data monitoring	21a	Data monitoring committee: all data monitors, delegated at participating centres, will constitute the data monitoring committee. Data recording A standard case report form (CRF) has been designed, to record in written all study details, and will be completed by the designated personnel. Any written change or correction must be dated, initialled and explained (if necessary) and must not obscure the original entry. In the patient's medical records, study participation, date of consent, assigned code and any other relevant information will be recorded. A check for data completeness and consistency will be performed weekly.
	21b	No reasons for study termination are identified. Patients and controls enrolled will be free to withdraw their consent
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct **N/A**
Auditing	23	External Project Advisory Committee (EPAC) of pain-OMICS FP7project will perform an overall scientific supervision of the trial and of emerging data.

Ethics and dissemination

Research ethics approval	24	The study protocol has been revised and approved (Local Ethical Committess) at all the clinical participating centres before initiating the enrolment.

Protocol amendments	25	The Coordinating Investigator and all participant Principal Investigators will act according to Article 10(a) of the Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001. Only amendments that are substantial will be notified to the Competent Authority (CA) and ethics committee concerned. In addition when an investigator must take urgent safety measures to protect the trial subjects from immediate hazard Article 10(b) allows them to do so before notifying the CA, but they must notify them as soon as possible. Non-substantial amendments The investigator does not have to notify non-substantial amendments to the documentation provided to the competent authority or the ethics committee. However, they should be recorded and if appropriate included in the next update of the Investigator Brochure and be available on request for inspection at the trial site. Substantial amendments Substantial amendments to the conduct of the clinical trial may arise from changes to the protocol or from new information relating to the scientific documents in support of the trial. Amendments to the trial will be regarded as “substantial” where they are likely to have a significant impact on: • the safety or physical or mental integrity of the subjects; • the scientific value of the trial; • the conduct or management of the trial; or • the quality or safety of any IMP used in the trial. In all cases, an amendment is only to be regarded as “substantial” when one or more of the above criteria are met. In the case the Coordinating Investigator will intend to make a substantial amendment to the protocol he will notify the concerned CA and relevant ethics committee, Substantial Amendment Form will be applied (Annex 2, “Detailed guidance for the request for authorisation of a clinical trial on a medicinal product for human use to the competent authorities, notification of substantial amendments and declaration of the end of the trial” October 2005) and procedures detailed in Article 10(a) of the Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 will be respected.

Consent or assent	26a	At each clinical unit a Delegation Log has been completed to provide a comprehensive list of study staff members and the duties that have been delegated to them by the Principal Investigator. Information Sheets specify, in an understandable language for patients, each aspect concerning their participation in the study (including protection, encoding and usage of data collected, moments for biological samples collection and biosamples uses). Moreover the Information Sheets and Informed Consent Form specify that the participation in the study is voluntary, provide patients with staff contact numbers (office number, e-mail or other) to be used in case of any doubts or question before the signature of Informed Consent or throughout the study, detail any aspects of the possible findings. Eligible patients will receive Information Sheets and to be included in the study patients have to sign an Informed Consent Form. Patients enrolled may always and without specification of reasons withdraw their informed consent. Informed Consent Form and Information Sheets are in language and terms understandable to the participants. Participants have the right: - To know that participation is voluntary; - To ask questions and receive understandable answers before making a decision; - To know the degree of risk and burden involved in participation; - To know if there are any benefits involved in participation; - To withdraw themselves and their biosamples from the project at any time; - To know how their data will be collected, protected during the project; Patients will receive a detailed description of the study purposes and planning during the enrolment visit by the local investigator (or a delegate), the visit will last about 30 minutes and all participation details and rights will be extensively described to the patient. Patients will have enough time to decide about their participation and any his/her question will be answered.
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable
Confidentiality	27	Once enrolled in the trial, patients will be assigned a unique anonymous code and data will be collected in the designated ad-hoc CRF. Each centre will keep separately the list of codes assigned to their patients and to samples collected. All study material will be maintained in a safe place site in each Pain Unit participating in the study, for 5 years. Clinical data are collected in the designated ad-hoc Case Report Form and into a dedicated web database (REDCap - Research Electronic Data Capture); access to the web-database is restricted to the project partners and can be accessed using a dedicated username and a password assigned to local principal investigators.

Declaration of interests	28	All the researchers participating to the study declare that they have no potential conflict of interest
Access to data	29	All details regarding data property, transfer and dissemination of foreground will be regulated by the Consortium Agreement signed by all the Parties involved in the Pain-OMICS project.
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation **N/A**
Dissemination policy	31a	The proposing group will manage patients' data and publications. The Scientific Board and Ethical Committee of PainOMICS Consortium will also review the results of the study in order to evaluate any possible societal impact of findings according to the ethical concerns about genetics/OMICS and diagnosis of chronic pain.
	31b	The authors of the publications will be decided on the basis of indications contained in the Uniform Requirements for Manuscripts (http://www.icmje.org/urm_full.pdf).
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code **N/A**
Appendices		
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

Items from the World Health Organization Trial Registration Data Set

Primary Registry and Trial Identifying Number
Clinicaltrials.gov (NCT02037789).

Date of Registration in Primary Registry
January 14, 2014

Source(s) of Monetary or Material Support
The present study is not funded by Industry or any other commercial sponsors.
This trial is a funded academic research; it is supported by founding from the European Commission in the context of the Seventh Framework Program of the European

Community for Research, Technological Development and Demonstration Activities. Grant agreement no: 602736.

Primary Sponsor
Azienda Ospedaliero-Universitaria di Parma, Italy

Secondary Sponsor(s)

Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
Multidisciplinary Pain Centre, Hospital Oost--Limburg (ZOL), Genk , Belgium
"St.Catharine" Orthopedics, Surgery, Neurology and Physical Medicine and Rehabilitation Specialty Hospital (St-Cat), Zabok, Croatia
The Center for Clinical Research (CPI), Winston-Salem, USA
Edith Cowan University (ECU), Perth, Australia
Genos Glycoscience Research Laboratory, Zagreb, Croatia
PolyOmica, Groningen, The Netherlands.
Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany.
Photeomix, IP Research Consulting SAS, Noisy le Grand, France
Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.

Contact for Public Queries
Massimo Allegri,MD. mallegri@parmanesthesia.com

Contact for Scientific Queries
Massimo Allegri,MD.
Department of Surgical Science, University of Parma, Via Gramsci 14, 43126 Parma, Italy
Phone: +39 0521 703567
E-mail: mallegri@parmanesthesia.com

Public Title
"OMICS" biomarkers associated with chronic low back pain: a retrospective clinical study

Scientific Title
"OMICS" biomarkers associated with chronic low back pain: a retrospective clinical study

Countries of Recruitment

Italy, Croatia, Belgium, USA, Australia

Health Condition(s) or Problem(s) Studied
Chronic Low back pain

Intervention(s)
N/A

Key Inclusion and Exclusion Criteria
Inclusion Criteria of patients with persistent CLBP:
age: older than 18;
chronic/persistent pain (pain lasting longer than 12 weeks) between the costal margins and gluteal fold, with or without symptoms into one or both legs
written informed consent signed;
Caucasian ancestry
Inclusion Criteria of healthy volunteers:
age: older than 18;
without any chronic/persistent pain (pain lasting longer than 12 weeks) in the last one year;
written informed consent signed;
Caucasian ancestry
Exclusion Criteria:
evidence of clinically unstable disease;
severe psychiatric disorder (excluding mild depression) or mental impairment;
recent history (< 1 year) of spinal fracture;
pain in the back due to spinal tumor or infection;
pregnancy

Study Type
Retrospective observational, multicenter, international clinical study, with a case control design

Date of First Enrollment
May, 2014

Target Sample Size
12000 individuals (9000 in the discovery phase and 3000 in the validation phase)
Divided in 4000 cases (CLBP) and 8000 controls (without CLBP).

Recruitment Status
Recruiting: participants are currently being recruited and enrolled

Primary Outcome(s)
GENETIC OUTCOME
The primary objective is to recognize genetic variants associated with persistent CLBP patients compared to patients without chronic/persistent pain. Through a Genetic Wide Association Study (GWAS) investigators will correlate genetic variants associated with persistent CLBP in a wide, international population of European ancestry.

Key Secondary Outcomes
GLYCOMIC AND ACTIVOMIC OUTCOME
Recognize Glycomic and Activomic data associated with persistent CLBP patients compared to patients without chronic/persistent pain. The sample size will better defined after the first interim analysis of first 400 patients.
STRATIFICATION OF OUR POPULATION

All "omic" data will be compared stratifying our population according to:
Pathophysiology: discogenic pain, spinal stenosis, facet joint pain, sacroiliac joint pain, low back pain with radicular pain (not predominant radicular pain) and widespread pain.
pain intensity
response to treatment
duration of pain

BMJ Open

"OMICS" BIOMARKERS ASSOCIATED WITH CHRONIC LOW BACK PAIN. PROTOCOL OF A RETROSPECTIVE LONGITUDINAL STUDY

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012070.R1
Article Type:	Protocol
Date Submitted by the Author:	15-Jul-2016
Complete List of Authors:	Allegri, Massimo; Universita degli Studi di Parma, Department of surgical science; Azienda Ospedaliera Universitaria Parma, Anesthesia Intensive Care and Pain Therapy service DE GREGORI, MANUELA; Fondazione IRCCS Policlinico San Matteo, Anesthesia, Intensive Care and Pain Therapy, Emergency Department MINELLA, CRISTINA; Fondazione IRCCS Policlinico San Matteo, Anesthesia, Intensive Care and Pain Therapy, Emergency Department Klersy, Catherine; Fondazione IRCCS Policlinico San Matteo, Service of Biometry & Satistics, Research Department Wang, Wei; Edith Cowan University, School of Medical and Health Sciences Sim, Moira; Edith Cowan University, School of Medical and Health Sciences Gieger, Christian; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit of Molecular Epidemiology; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Institute of Epidemiology II Manz, Judith; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit of Molecular Epidemiology; Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Institute of Epidemiology II Pemberton, Iain; Photeomix, IP Research Consulting SAS MacDougall, Jane; Photeomix, IP Research Consulting SAS Williams, Frances; Kings College London, Twin Research and Genetic Epidemiology Depart Van Zundert, Jan; Ziekenhuis Oost-Limburg, Department of Anesthesiology, Critical Care and Multidisciplinary Pain Center Buyse, Klaas; Ziekenhuis Oost-Limburg, Department of Anesthesiology, Critical Care and Multidisciplinary Pain Center Gudelj, Ivan; Genos Glycoscience Research Laboratory Lauc, Gordan; Genos Glycoscience Research Laboratory Primorac, Dragan; St. Catherine Specialty Hospital; Sveuciliste u Splitu Medicinski fakultet Skelin, Andrea; Genos Glycoscience Research Laboratory; St. Catherine Specialty Hospital Aulchenko, Yurii; PolyOmica Karssen, Lennart; PolyOmica Kapural, Leonardo; Carolinas Pain Institute Rauck, Richard; Carolinas Pain Institute Fanelli, Guido; Universita degli Studi di Parma, Department of surgical science; Azienda Ospedaliero Universitaria di Parma, Anesthesia Intensive Care and Pain Therapy service

Primary Subject Heading:	Patient-centred medicine
Secondary Subject Heading:	Genetics and genomics, Research methods
Keywords:	Pain management < ANAESTHETICS, CHRONIC LOW BACK PAIN, GENOME-WIDE ASSOCIATION STUDY, GLYCOMICS, ACTIVOMICS

"OMICS" BIOMARKERS ASSOCIATED WITH CHRONIC LOW BACK PAIN. PROTOCOL OF A RETROSPECTIVE LONGITUDINAL STUDY

Authors: Massimo Allegri[1,2], Manuela De Gregori[3], Cristina E. Minella[3], Catherine Klersy[4], Wei Wang[5], Moira Sim[5], Christian Gieger[6,7], Judith Manz[6,7], Iain K. Pemberton[8], Jane MacDougall[8], Frances MK Williams[9], Jan Van Zundert[10], Klaas Buyse[10], Gordan Lauc[11], Ivan Gudelj[11], Dragan Primorac[12,13,14,15], Andrea Skelin[11,12], Yurii S. Aulchenko[17], Lennart C. Karssen[17], Leonardo Kapural[18], Richard Rauck[18], Guido Fanelli[1,2] on behalf of the PainOMICS Group*

PainOMICS Group* Collaborators

Melanie Waldenberger, PhD[6,7]

Manuela Zunino, MD[3]

Alice Montalti, MD[3]

Dario Bugada, MD[2]

Concetta Dagostino, PhD[2]

Silvana Montella, MD[2]

Marco Baciarello, MD[2]

Christian Compagnone, MD[2]

1 Department of surgical science. University of Parma, Parma, Italy.

2 Anesthesia Intensive Care and Pain Therapy service. Azienda Ospedaliera Universitaria Parma, Parma, Italy.

3 Anesthesia, Intensive Care and Pain Therapy, Emergency Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

4 Service of Biometry & Satistics, Research Department, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

5 School of Medical and Health Sciences, Edith Cowan University

6 Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany.

7 Institute of Epidemiology II, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany.

8 Photeomix, IP Research Consulting SAS, Noisy le Grand, France.

9 Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.

10 Department of Anesthesiology, Critical Care and Multidisciplinary Pain Center, Ziekenhuis Oost-Limburg, Genk, Belgium.

11 Genos Glycoscience Research Laboratory, Zagreb, Croatia

12 St. Catherine Specialty Hospital, Zabok, Croatia

13 Penn State University, Eberly College of Science, State College, PA, USA

14 Faculty of Medicine University of Osijek, Osijek, Croatia

15 University of Split School of Medicine, Split, Croatia

16 Children's Hospital Srebrnjak, Zagreb, Croatia

17 PolyOmica, Groningen, The Netherlands.

18 Carolinas Pain Institute, Winston-Salem, NC 27103, USA.

Corresponding Author:

Dr Massimo Allegri
Anesthesia Intensive Care and Pain Therapy service, Azienda Ospedaliera Universitaria Parma
via Gramsci 14, 43126 Parma (ITALY)
phone: 00390521703965
fax: 00390521703567
email: mallegri@parmanesthesia.com

Abstract

Introduction: Chronic low back pain (CLBP) produces considerable direct costs as well as indirect burdens for society, industry and health systems.

CLBP is characterized by heterogeneity, inclusion of several pain syndromes, different underlying molecular pathologies, and interaction with psychosocial factors that leads to a range of clinical manifestations. There is still much to understand in the underlying pathological processes and the non-psychosocial factors, which account for differences in outcomes.

Biomarkers that may be objectively used for diagnosis and personalized, targeted and cost-effective treatment are still lacking. Therefore, any data that may be obtained at the "-omics" level (glycomics, Activomics and genome-wide association studies - GWAS) may be helpful to use as dynamic biomarkers for elucidating CLBP pathogenesis and may ultimately provide prognostic information too.

By means of a retrospective, observational, case-cohort, multicenter study, we aim to investigate new promising biomarkers potentially able to solve some of the issues related to CLBP.

Methods and analysis: the study follows a two-phase, 1:2 case-control model. A total of 12000 individuals (4000 *cases* and 8000 *controls*) will be enrolled; clinical data will be registered, with particular attention to pain characteristics and outcomes of pain treatments. Blood samples will be collected to perform omics studies. The primary objective is to recognize genetic variants associated with CLBP; secondary objectives are to study glycomics and Activomics profiles associated with CLBP.

Ethics and dissemination: The study is part of the PainOMICS project funded by European Community in the 7th Framework Program. The study has been approved from competent Ethical Bodies and copies of approvals were provided to the European Commission before starting the study.

Results of the study will be reviewed by the Scientific Board and Ethical Committee of the PainOMICS Consortium.

The scientific results will be disseminated through peer-reviewed journals.

Registration details: Registered on Clinicaltrials.gov (NCT02037789).

Strengths and limitations of this study

Strength:

- Multiple-centre and multiple-discipline study: the study includes centres in the EU, USA and Australia, with research teams specialising in: (1) clinical aspects of pain, (2) biology and genetics of pain, (3) generation of 'omics' data, and (4) analysis of multiple 'omics' data.
- Hypothesis driven v.s. hypostasis generating: The study aims to profile "-omics" biomarkers (GWAS, glycomics and activomics) potentially to decipher the pathogenesis of CLBP associated with the different patho-physiological patterns.
- Longitudinal design with a large sample size: 1) Discovery phase 3000 cases and 6000 controls; 2) validation phase: 1000 cases and 2000 controls.

Limitations:

- While the heterogeneity of the study populations is helpful in the discovery phase, this may limit conclusions in the validation phase.
- Functional investigation with animal model has not been included in the current project due to the limitation of the funding.

Introduction

Low Back Pain (LBP) is one of the most common health problems worldwide with an estimated age-standardized point prevalence of 9.4%[1]. In 2012 a global review of the prevalence of low back pain was published reporting a mean ± SEM point prevalence of activity-limiting low back pain lasting more than 1 day of 11.9 ± 2.0%, and the 1-month prevalence of 23.2 ±2.9%[2]. LBP accounts for considerable disability and work absence, and ranks in the Global Burden of Disease 2013 study as a leading contributor to global disability measured in years lost due to disability (YLDs)[3]. LBP is defined as pain and discomfort, localized below the costal margin and above the inferior gluteal folds, with or without leg pain.

Prognostic factors for LBP include demographic factors (educational attainment, age, gender), occupational factors (employment), mental health morbidity (anxiety, depression), perception of pain and disability (pain intensity and expectation of persistent pain) and other psychological factors (fear avoidance, catastrophizing, illness perceptions) [4].

LBP becomes chronic low back pain (CLBP) when symptoms last at least 3 months. Activity-limiting LBP tends to recur and the course of LBP is increasingly viewed as a chronic recurring condition[5], which accounts for considerable direct economic costs as well as indirect burdens for society, industry and health systems[6, 7]. The prevalence of chronic low back pain (CLBP) appears to be rising with an increase from 3.9% in 1992 to 10.2% in 2006 in the USA [8], and from 4.2% in 2002 to 9.6% in 2010 in Brazil[9].

Pain is a subjective sensation, which is influenced by a range of physical and psychosocial factors through poorly understood neural mechanisms[10]. The advances in medical imaging have improved our ability to identify the anatomic origin of CLBP, however CLBP is heterogeneous and the anatomic site only tells part of the story. There is considerable variation in prognosis and, at present, CLBP exerts a substantial burden on the individual, the family and workplace.

Evidence based guidelines recommend the initial exclusion of serious diagnosis before the implementation of clinical management, which promotes continued function and best practice rehabilitation approaches. However, there is much more to understand about the underlying process, how this affects the prognosis and how we can use this to tailor treatment for the individual. In making a diagnosis of LBP, red flag symptoms are used to identify the need for investigation for underlying serious illness[11]. Yellow flags (psychosocial factors) identify risks of chronicity [12]and highlight the heterogeneity and complexity of CLBP where the severity, chronicity and prognosis may depend on the anatomical site, the underlying pathological process, comorbidities as well as individual psychosocial factors.

While psychosocial factors clearly influence outcomes in CLBP, genetic and epigenetic factors may account for some variation in response to treatment. Even though both persistent CLBP and disc degeneration are known to be heritable[13, 14] and the two traits are highly related to one another, with disc degeneration being a major predictor for LBP episodes [14], few genetic variants have been identified and confirmed for both traits [15-17]. Only two genome wide association studies report on chronic/persistent widespread pain [16, 17] and two GWAS of intervertebral disc degeneration [18, 19]. In keeping with other common complex traits, the individual effects of the identified loci are small and explain only a small fraction of the trait or disease variation [20]. As such, they do not substantially improve predictions over those based on known factors such as family history [21-23]. Unfortunately these data have not yet shed light on the pain pathogenesis mechanisms and they do not offer prediction of treatment likely suitable in individual patients. Replication of these findings is also needed. Hence, despite promising recent data, new studies are needed to identify objective biomarkers for both diagnosis and prediction of treatment's efficacy in CLBP patients.

Glycomics is an emerging field, recently identified as a priority for the next decade by the US National Academies of Science [24, 25]. Recent studies reported protein glycosylation in large human population samples, with promising glycan profiling for disease diagnosis and stratification e.g., autoimmune diseases and hematological cancers, metabolic syndrome, systemic lupus erythematosus and many other diseases [26-33]. Together with glycomics, Activomics may also provide insight via new biomarkers for low back pain as it combines data involved in enzymatic activity of several post-translational modification proteins in an integrated model, providing dynamic characterization of the current state of an organism.

Activomics, is a novel –omic strategy that aims to describe biological systems in terms of differential protein post-translational modification (PTM) activities. Perturbations in intracellular and/or intercellular cell signalling networks are frequently linked to chronic diseases such as cancer. Enzyme activities in serum are monitored using a proprietary panel of protein and peptide substrates under multiple assay conditions including the judicious use of enzyme cofactors and inhibitors in order to optimize the discrimination

between protein modification enzymes with preferences for overlapping primary sequence or structural targets. Principal Activomics substrate panels include proteases (metalloproteinases, serine, cysteine, aspartic proteases) and their protease inhibitors (e.g. serpins), caspases, kinases (ser/thr and tyr), phosphatases and (de)acetylases.

Here we present the study protocol for a retrospective analysis, in a large cohort of CLBP patients, to determine "-omics" biomarkers (GWAS, glycomics and Activomics) potentially associated with susceptibility to CLBP and with different patho-physiological patterns [34, 35]. Both glycomic and activomic approaches aim to reveal alterations in proteome complexity that arise from post translational modifications that vary in response to changes in the physiological environment, a particularly important avenue to explore in chronic inflammatory diseases. Furthermore, exploring disease-related links between glycomic and activomic data within the context of a clearly defined genetic and demographic background is a highly original and potentially instructive secondary objective of the current study. Since aforementioned studies connected mostly N-glycans with chronic inflammation and methods for high-throughput glycoprotein analysis are still in development, our glycomic data will be based exclusively on released N-glycans.

This study will link clinical data to the multiple "-omics" analyses, thereby profiling novel biomarkers, which have strong potential to advance our knowledge of some of the remaining unsolved problems in CLBP.

The present manuscript serves to describe the registered protocol in order to disseminate the rationale, the methods and the main aims of the clinical study.

Methods

We present a retrospective observational, multicenter, international clinical study, with a case control design. We describe the details of the retrospective cohort protocol clinical study without providing any preliminary results. Patient enrolment is currently active, up to September 2016. The study is part of the PainOMICS project that includes four different trials and was reviewed and funded by European Community in the 7th Framework Program (FP7) - THEME [HEALTH.2013.2.2.1-5 - Understanding and controlling pain].

The project includes six clinical centers from Italy, Croatia, Belgium, Australia and the United States and four centers for scientific analyses in Croatia, France, Germany and the United Kingdom. Statistical expertise is provided by the "PolyOmica" consulting based in the Netherlands.

The retrospective study was approved by the ethical committees of each separate clinical center between December 2013 and March 2014.

The study is registered at Clinicaltrials.gov (NCT02037789).

Participant enrolment and data collection

Cases (patients having CLBP) will be enrolled by each participating center (Figure 1). Every effort will be made to accumulate a well characterized cohort of patients with persistent CLBP, sub-grouped according to the likely anatomical cause of the pain. Patients fulfilling the following conditions will be considered for enrolment: age 18 years and older; chronic pain (pain lasting longer than 12 weeks) between the costal margins and gluteal fold, with or without symptoms in one or both legs, written informed consent signed and Caucasian ancestry.

Controls (patients without chronic low back pain) will be retrieved from two different sources: 1) existing bio-banks of healthy subjects having collected information about CLBP; 2) subjects enrolled in the parallel prospective study on acute LBP (part of PainOMICS project – NCT 02037763), i.e. patients who presented with acute LBP and have not become chronic over 6 months but the pain has resolved. Age (decades) and gender distribution of controls will be matched as closely as possible to that of the cases. Controls are enrolled according to the following inclusion criteria: older than 18 years, no chronic pain (lasting longer than 12 weeks) in the past 12 months, written informed consent obtained and Caucasian ancestry.

Subjects with any evidence of clinically unstable disease, severe psychiatric disorder (excluding mild depression) or mental impairment; recent history (less than one year) of spinal fracture, back pain due to spinal tumor or infection, pregnancy, will be excluded from the study.

Patients and controls selected for participation will receive a detailed description of the study and will be asked to sign an informed consent prior to entering the study (Enrolment Visit).

Once enrolled in the trial, patients will be assigned to a unique anonymous code. Data collection includes; demographics (age, gender, race, body mass index - BMI, occupational history), clinical and pharmacological history, pain characteristics (onset, duration, intensity, pain referral pattern, irradiation, sensory abnormalities, precipitating events, history of previous episodes), effectiveness/tolerability of pain treatments received (when applicable). A specific questionnaire (Pain-DETECT, PD) is applied to evaluate both pain type and the pain generator, as well as the possible pathophysiological (nociceptive and/or neuropathic) mechanism sustaining CLBP and functional impairment.

Blood samples for omics analyses are taken from each enrolled patient at the time of the enrolment consultation, and biological samples are sent to analytical partners of the Consortium for specific omics analysis.

Clinical data are collected in the designated ad-hoc Case Report Form and into a dedicated web database (REDCap - Research Electronic Data Capture); access to the web-database is restricted to the project partners and can be accessed using a dedicated username and a password.

Omics data are centralized in a specific supervised database.

Samples collection methodology

All the clinical centers have to guarantee that the biological samples from each patient will be prepared, stored, and shipped following the analytical procedures described in three standard operating procedures

(SOPs) developed and validated to provide details for conducting the study, phase by phase, with written instructions to achieve uniformity of the procedures used for obtaining patient blood for omic analysis techniques, storing the samples, and shipping the aliquots to the specialized laboratories.

All the patients and controls undergo blood sampling for omics determinations. The samples will be collected into two tubes containing EDTA for genetic and glycomic analyses and into one serum tube with clot activator plus gel for the Activomics. For genetic analyses, DNA will be isolated from whole blood samples using a commercial DNA extraction kit.

Detailed description of the validation of SOPs is provided in a separate paper under submission.

Analyses for biologic markers

Genetic analyses:

GWAS analyses will be performed on DNA samples isolated from whole blood, using genome-wide Illumina genotyping technology [36, 37]. Briefly, 4µl of 60 ng/µl DNA will be amplified at 37 °C overnight, followed by enzymatic fragmentation, alcohol precipitation and DNA resuspension. Whole genome amplified DNA will be hybridized to the Illumina HumanCore Bead Chip, including > 240,000 genome-wide tag SNPs and > 20,000 high-value markers (indels and updated exome-focused content). After hybridization, allelic specificity will be conferred by enzymatic base extension. Products will be subsequently stained and the intensities of the beads' fluorescence will be detected by the iScan system (Illumina, Inc.) Genotype calling will be performed with the Illumina GenomeStudio 2011.1 Genotyping Module 1.9.4 software. During quality control genotype data will be filtered by sample-wise and variant-wise call rates and by Hardy-Weinberg Eqilibrium p-values. Related individuals, individuals with extreme heterozygosity rates, and individuals whose genetics suggest non-Caucasian origins will be discarded. A much denser set of markers will be obtained by genotype imputation using the Haplotype Reference Consortium (HRC) reference panel. The resulting set of markers will contain both genotypes measured on the Illumina chip and genotypes imputed based on the very dense HRC reference panel.

For each genetic variant a standard association model (linear or logistic regression model) will be used to investigate associations between CLBP-related phenotypes and genetic markers. The multiple testing problem will be addressed by judging the significance of associations using a Bonferroni-corrected, genome-wide significance level corresponding to a nominal significance level of 5 %.

Glycomics analyses:

Glycomics analyses will be performed on total serum proteins and on a single serum protein, immunoglobulin G (IgG). IgG will be isolated from serum samples by affinity chromatography using 96-well monolithic plates with Protein G as previously described (Pucic et al, MCP, 2011). N-glycans will be released from total serum proteins and IgG by overnight deglycosylation with N-glycosidase F (PNGase F). Released N-glycans will be fluorescently labeled with 2-aminobenzamide (2-AB) fluorescent tag and purified by hydrophilic interaction liquid chromatography (HILIC) solid phase extraction (SPE). Labeled N-glycans will be analyzed by hydrophilic interaction chromatography on a Waters Acquity UPLC instrument

using Waters BEH Glycan chromatography column, 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. The system will be calibrated using an external standard of hydrolyzed and 2-AB labeled glucose oligomers from which the retention times for the individual glycans will be converted to glucose units. Data processing will be performed with an automatic processing method after which each chromatogram will be manually corrected to maintain the same intervals of integration for all the samples. The chromatograms obtained will all be separated in the same manner into peaks and the amount of glycans in each peak will be expressed as percentage of total integrated area.

Activomics analyses:

Activomics analyses will be performed on retrospective serum samples from anonymous but well characterized patients. Samples are collected, handled and analyzed in a way that minimizes freeze-thaw cycles and arrayed with the aid of a automatic liquid handler robot (Multiprobe II, Perkin Elmer) in 96-well microtitre plates for high throughput screening using microfluidic mobility shift assays (insert reference Drueckes P). For each enzymatic reaction tested, 2 microliters of serum from patients and healthy controls will be incubated with the appropriate Activomics substrate under controlled conditions (time, temperature, optimized buffer conditions, etc.). In general, the fluorescent peptides to be used as target substrates are synthesized through the addition of an N-terminal carboxyfluorescein (FAM) group via an aminohexanoic acid (Ahx) spacer group, i.e. FAM-Ahx-Peptide. High throughput screening is performed using a modification of the capillary electrophoretic mobility shift assay on an automated microfluidic platform (EZReader, Perkin Elmer, USA). Peptide substrates are generally C-terminally amidated and purified by HPLC to > 90% purity (Genscript, Hong Kong). All lyophilized peptides are redissolved in sterile, double distilled water at 2 mM concentration and are diluted to 10 µM for screening tests. Assays are performed in the appropriate reaction buffer in 384-well format (Corning 3821 BC) using a semi-automated pipettor system for reproducibility (Sorensen Multi, precision CV < 5%). Post-translationally modified products are separated from unreacted substrate by high voltage microfluidic mobility shift assays using voltage and pressure parameters optimized for each substrate. Product conversion is assessed from the respective peak areas obtained from electrophoretic mobility shifts (% product/substrate + product). The extent of PTM of each substrate is assessed and compared in a univariate analysis for patients with chronic versus those without low back pain. . The assays will be repeated for a panel of different substrates in order to provide a wide view of disease-related changes to post-translational modification activities for multivariate statistical analysis. Performance characteristics of the panel will be assessed by univariate and multivariate hierarchical clustering and principal component analysis to differentiate activities of disease versus control samples. Sensitivity and specificity will be evaluated in ROC analyses to define cut-off values in the design of the optimal predictive biomarker panel.

Primary and secondary objectives

The primary objective of this retrospective study is to recognize genetic variants associated with persistent CLBP, by comparing CLBP patients and pain-free patients. We will correlate genetic variants associated with CLBP through a GWAS study, in a wide international population of European ancestry.
Secondary objectives are to recognize glycomic and Activomic data associated with CLBP patients compared to patients without CLBP.

The participating clinical centers have defined a minimal shared diagnostic dataset available in all clinical centers. Each of the participants will stratify patients according to their clinical features, imaging data and results from Pain-DETECT. Considering the patient's response to diagnostic procedures, patients will be sub-grouped (taking into account the patient history, clinical examination, radiological results and potentially the response to diagnostic blocks) into 6 main categories:

1. spinal stenosis,
2. discogenic pain,
3. facet joint pain,
4. sacroiliac joint pain,
5. low back pain with radicular pain (not predominant radicular pain),
6. widespread low back pain.

Statistics

Study design

The study will follow a two-phase (discovery & validation), 1:2 case-control model:

- Discovery population: random sample of two thirds of the entire population of cases.
- Validation population: the remaining one third. Following the GWAS phase, the genes discovered will be assessed for biological plausibility and entered into the validation phase

Sample size calculation

Since, to date, no data on the omics of CLBP are available in literature, and since the variants associated with most common diseases have modest effects we considered a number of scenarios, ranging in model assumptions with respect to allele frequency and effect.
With 3000 cases and 6000 controls, we assessed genetic scenarios in which we have 80% power to detect association at genome-wide significant level[38]. Consistent with the literature, for high odds ratios (OR=2) we will be able to detect variants with low minor allele frequency (MAF≥1.5%). With higher allele frequency we will be able to find smaller effects; for example, for variants with MAF=25% we have power to detect OR≥1.25

For the replication/validation phase, with 1000 cases and 2000 controls, we have 80% power to confirm detected variants when using nominal p = 0.005 (leading to experiment-wise type I error of 0.05 assuming 10 tests) and approximately 60% power in case of more severe multiple testing (100 tests, leading to nominal p = 0.0005).

Statistical analysis

All the enrolled patients and controls will be analyzed.

Discovery phase: For analyses investigating highly dimensional omics space we will use a range of approaches. The first approach will extend the classical sequential framework. Predictor screening will be performed by logistic or Cox regression models traversing through the omics space and incorporating few predictors at a time. Statistically significant (at experiment-wise level) predictors will be included in the model, and the next iteration through the omics space will be performed. A classical example of this approach includes genome-wide association analysis, followed by conditional analyses for identification of secondary signals. To investigate a large numbers of predictors simultaneously, we will use modern regularization/shrinkage and machine learning methods allowing analysis of (relatively) large numbers of predictors jointly. While this type of approach does not address the question of statistical testing in the same way as "classical" approaches do, it is widely used in the context of biomarker discovery, where prediction and not the *p*-values are of primary interest. For all methods aimed towards biomarker discovery, the accuracy of prediction will be accessed by cross-validation, and optimal solutions will be analyzed to identify potential biomarkers, which will be selected on their discriminative value in a receiver operator characteristic (ROC) analysis.

Validation phase: The discovered genetic variants will be examined bio-informatically for biological plausibility before entering the validation phase. The association of the candidate polymorphisms with the outcome (being a case) will be assessed with logistic regression. The following strategies will be used: single genes assessment/genetic score (sum of candidate genes)/multiple genes. Adjustment for covariates (age, gender, clinical features) will be performed.

Details of the statistical analyses will be provided in the final statistical analysis plan (SAP).

The analysis of the secondary endpoints will follow the same principles reported above.

Ethical Issues

Sample collection and use of clinical data has started only after the Ethical Approval of the present study protocol from the competent ethical bodies (Ethics Committees of the Institutions involved in patients enrolment).

Copies of Ethical Approval were provided to the European Commission before initiating the study. All protocol, copies of Informed Consent and Information Sheets approved by the competent ethical bodies, were provided to the European Commission before starting the study.

The Scientific Board and Ethical Committee of PainOMICS Consortium will also review the results of the study in order to evaluate any possible societal impact of our findings according to the ethical concerns about genetics/OMICS and diagnosis of chronic pain [39].

Monitoring and quality assessment

Patients will be withdrawn from the study in case of withdrawal of consent (subjects may always and without obligation withdraw their informed consent), or any other condition that, upon clinical judgment of the investigator, will make unacceptable further study participation for that individual patient.

The Coordinating Investigator (University Hospital of Parma, Italy) will delegate, in each participating center, a clinical supervisor (to ensure that the study is conducted according to the protocol, to good clinical practice, and to national regulations) and also a data monitor, to ensure accuracy, completeness and verification of patients' data. The data monitors, from each participating center, will make up the data monitoring committee. The External Project Advisory Committee (EPAC) of the Pain-OMICS FP7 project will perform an overall scientific supervision of the trial and of the emergent data.

The participating members will discuss results and any issues of the study at regular audits during the annual SIMPAR (Study in Multidisciplinary Pain Research) meeting, and in any other case that may be deemed necessary.

Conclusions

This study is, to the best of our knowledge, the first to investigate genetic and OMIC biomarkers in a large population sample of CLPB patients. These biomarkers may be related to pain sensation, as well as to disease pathophysiology and pain generators. The overall objective is to validate associations that may result in a more personalized diagnosis and therapy of a disease with a high health and societal burden such LBP.

Furthermore, the novel biomarkers emerging from this retrospective study will be validated in a prospective cohort collected within the same PainOMICS project, in order to evaluate their ability to predict the possibility of advancement to chronic pain in patients suffering from an acute episode of LBP.

A possible bias could be also related to the fact that in some patients pain could be still related to acute inflammation even though pain was lasting since more three months. However, as we enroll all patients evaluated in chronic pain services who were referred after several pharmacological therapies, we think that this bias will be limited also by the high number of patients enrolled.

The Pain-OMICS project is expected to significantly expand the level of knowledge on how low back pain is generated, propagated and quenched. We will mobilize significant human and material resources in Europe and USA, allowing a comprehensive characterization of large cohorts of CLPB patients, aiming to identify a number of potential biomarkers related to different aspects of chronic low back pain, as well as potential new targets for therapy.

With this protocol we would like to investigate better biomarkers related to chronic low back pain". The next mandatory step will be to evaluate if and how these biomarkers could help to predict patients at higher risk of developing chronic pain after an acute episode and how these biomarkers might also be related to predicting response to pharmacological/surgical treatment. The same research group is already conducting a new prospective study investigating the transition from acute to chronic low back pain and the enrollment will be closed in the spring 2017.

Authors' contributions:

MA, CG, IKP, FMKW, JVZ, GL, DP, YSA, LK and GF conceived the study and revised the paper.

MA, CEM drafted the paper. MDG, CK, WW, MS, JM, JMD, KB, IG, AS, LCK, RR provided feedback on the manuscript, and all authors reviewed and approved the final version of the paper.

Acknowledgements

MW and CD provided scientific support and are involved in sample management and analysis.

MZ, AM, DB, SM, MB, CC are collecting data and caring for study patients.

Funding statement

This trial is funded academic/SME research; it is supported by funding from the European Commission in the context of the Seventh Framework Program of the European Community for Research, Technological Development and Demonstration Activities - (FP7) - THEME [HEALTH.2013.2.2.1-5 - Understanding and controlling pain].

The present study is not funded by Industry or any other commercial sponsors.

Competing Interests

The study is supported by a grant from the European Commission (602736).

Dr Allegri is a consultant for Grunenthal, Angelini and Mundipharma. He also collaborated for speeches with MSD and Carefusion.

Dr. Lauc has multiple patents in the field of glycoscience issued.

Dr Iain K Pemberton is the research director of Photeomix, the commercial name of IP Research Consulting SAS, who have ongoing patent applications and trademarks related to Activomics.

Dr. Aulchenko is a director and co-owner of Maatschap PolyOmica, which provides (consulting) services in the area of (statistical) (gen)omics.

Data sharing statement

No additional data are available

References

1. Hoy, D., et al., *The global burden of low back pain: estimates from the Global Burden of Disease 2010 study.* Ann Rheum Dis, 2014. **73**(6): p. 968-74.

2. Hoy, D., et al., *A systematic review of the global prevalence of low back pain.* Arthritis Rheum, 2012. **64**(6): p. 2028-37.
3. Collaborators, G.B.o.D.S., *Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.* Lancet, 2015. **386**(9995): p. 743-800.
4. Campbell, P., et al., *Prognostic indicators of low back pain in primary care: five-year prospective study.* J Pain, 2013. **14**(8): p. 873-83.
5. Hoy, D., et al., *The Epidemiology of low back pain.* Best Pract Res Clin Rheumatol, 2010. **24**(6): p. 769-81.
6. Steenstra, I.A., et al., *Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature.* Occup Environ Med, 2005. **62**(12): p. 851-60.
7. Organization, W.H., *Priority medicines for Europe and the worldupdate report 2013. Chapter 6.*
8. Freburger, J.K., et al., *The rising prevalence of chronic low back pain.* Arch Intern Med, 2009. **169**(3): p. 251-8.
9. Meucci, R.D., et al., *Increase of chronic low back pain prevalence in a medium-sized city of southern Brazil.* BMC Musculoskelet Disord, 2013. **14**: p. 155.
10. Koyama, T., et al., *The subjective experience of pain: where expectations become reality.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12950-5.
11. http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/red-flags-in-back-pain. Available from: http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/red-flags-in-back-pain.
12. http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/yellow-flags-in-back-pain. Available from: 12. http://www.sheffieldbackpain.com/professional-resources/learning/in-detail/yellow-flags-in-back-pain.
13. Battie, M.C., et al., *Heritability of low back pain and the role of disc degeneration.* Pain, 2007. **131**(3): p. 272-80.
14. MacGregor, A.J., et al., *Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins.* Arthritis Rheum, 2004. **51**(2): p. 160-7.
15. Eskola, P.J., et al., *Genetic association studies in lumbar disc degeneration: a systematic review.* PLoS One, 2012. **7**(11): p. e49995.
16. Kim, H., D. Clark, and R.A. Dionne, *Genetic contributions to clinical pain and analgesia: avoiding pitfalls in genetic research.* J Pain, 2009. **10**(7): p. 663-93.
17. Peters, M.J., et al., *Genome-wide association study meta-analysis of chronic widespread pain: evidence for involvement of the 5p15.2 region.* Ann Rheum Dis, 2013. **72**(3): p. 427-36.
18. Williams, F.M., et al., *Novel genetic variants associated with lumbar disc degeneration in northern Europeans: a meta-analysis of 4600 subjects.* Ann Rheum Dis, 2013. **72**(7): p. 1141-8.
19. Song, Y.Q., et al., *Lumbar disc degeneration is linked to a carbohydrate sulfotransferase 3 variant.* J Clin Invest, 2013. **123**(11): p. 4909-17.
20. Visscher, P.M., et al., *Five years of GWAS discovery.* Am J Hum Genet, 2012. **90**(1): p. 7-24.
21. Aulchenko, Y.S., et al., *Predicting human height by Victorian and genomic methods.* Eur J Hum Genet, 2009. **17**(8): p. 1070-5.
22. Clayton, D.G., *Prediction and interaction in complex disease genetics: experience in type 1 diabetes.* PLoS Genet, 2009. **5**(7): p. e1000540.
23. Eichler, E.E., et al., *Missing heritability and strategies for finding the underlying causes of complex disease.* Nat Rev Genet, 2010. **11**(6): p. 446-50.

24. *National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. Transforming Glycoscience: A Roadmap for the Future*. 2012; available on http://www.ncbi.nlm.nih.gov/books/NBK109958/, Washington (DC): National Academies Press (US).
25. Lauc, G., et al., *Complex genetic regulation of protein glycosylation.* Mol Biosyst, 2010. **6**(2): p. 329-35.
26. Knezevic, A., et al., *Variability, heritability and environmental determinants of human plasma N-glycome.* J Proteome Res, 2009. **8**(2): p. 694-701.
27. Lauc, G., et al., *Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers.* PLoS Genet, 2013. **9**(1): p. e1003225.
28. Lauc, G., et al., *Mechanisms of disease: The human N-glycome.* Biochim Biophys Acta, 2015.
29. Lu, J.P., et al., *Screening novel biomarkers for metabolic syndrome by profiling human plasma N-glycans in Chinese Han and Croatian populations.* J Proteome Res, 2011. **10**(11): p. 4959-69.
30. Pabst, M., et al., *IL-1beta and TNF-alpha alter the glycophenotype of primary human chondrocytes in vitro.* Carbohydr Res, 2010. **345**(10): p. 1389-93.
31. Pucic, M., et al., *High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations.* Mol Cell Proteomics, 2011. **10**(10): p. M111 010090.
32. Ruhaak, L.R., et al., *Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity.* PLoS One, 2010. **5**(9): p. e12566.
33. Vuckovic, F., et al., *Systemic lupus erythematosus associates with the decreased immunosuppressive potential of the IgG glycome.* Arthritis Rheumatol, 2015.
34. Freidin, M.B., et al., *Using omics in chronic pain conditions to delineate mechanisms and provide new therapeutic strategies.* Pain Manag, 2016. **6**(3): p. 211-5.
35. Freidin, M.B., et al., *The Association Between Low Back Pain and Composition of IgG Glycome.* Sci Rep, 2016. 6:26815.
36. Scott, R.A., et al., *Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways.* Nat Genet, 2012. **44**(9): p. 991-1005.
37. Schaefer, A.S., et al., *Genetic evidence for PLASMINOGEN as a shared genetic risk factor of coronary artery disease and periodontitis*. Circ Cardiovasc Genet, 2015. **8**(1): p.159-67.
38. Pe'er, I., et al., *Estimation of the multiple testing burden for genomewide association studies of nearly all common variants.* Genet Epidemiol, 2008. **32**(4): p. 381-5.
39. Compagnone, C., et al., *Ethical issues in pain and omics research. Some points to start the debate.* Croat Med J, 2014. **55**(1): p. 1-2.

Figure 1

Study flow chart

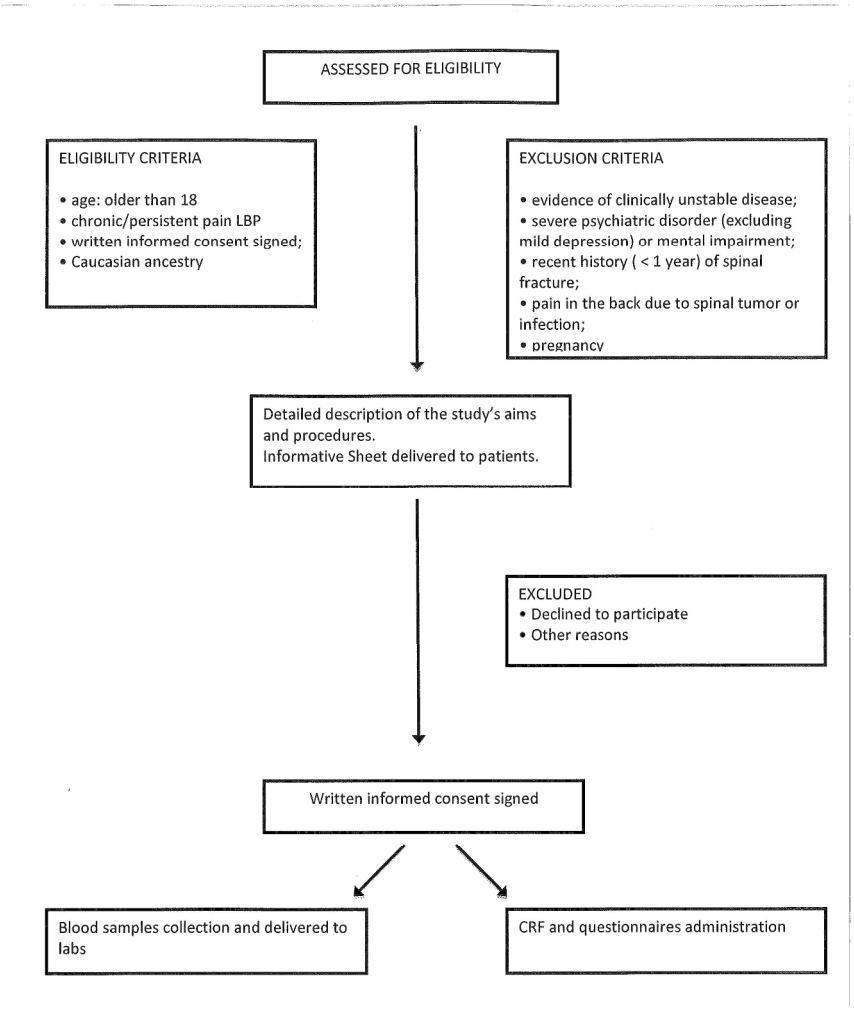

Figure 1
Study flow chart

The legend is on page 14
209x297mm (300 x 300 DPI)